

methoden daten analysen

ZEITSCHRIFT FÜR EMPIRISCHE SOZIALFORSCHUNG

mda

2008, Jahrgang 2, Heft 2



Anna Schnauber und Gregor Daschmann Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews?

Christian Weidmann, Patrick Schmich und Sabine Schiller-Born Kontrollüberzeugungen der Interviewer und Teilnahme an telefonischen Befragungen

Christian Babka von Gostomski und Monika Pupeter Zufallsbefragung von Ausländern auf Basis des Ausländerzentralregisters

Martin Kroh, Rainer Pischner, Martin Spieß und Gert G. Wagner Non-Original Sample Members in the German Household Panel Study (SOEP)

Herausgegeben von *Christof Wolf
Marek Fuchs
Bärbel Knäuper
Petra Stein*

Methoden – Daten – Analysen. Zeitschrift für Empirische Sozialforschung

Die Zeitschrift wird herausgegeben von GESIS – Leibniz-Institut für Sozialwissenschaften.

Herausgeber: Christof **Wolf** (Mannheim, geschäftsführend), Marek **Fuchs** (Kassel), Bärbel **Knäuper** (Montreal), Petra **Stein** (Duisburg-Essen)

Wissenschaftlicher

Beirat: Hans-Jürgen **Andreß** (Köln), Andreas **Diekmann** (Zürich), Sabine **Häder** (Mannheim), Udo **Kelle** (Marburg), Dagmar **Krebs** (Gießen), Frauke **Kreuter** (College Park, Maryland), Edith **de Leeuw** (Utrecht), Norbert **Schwarz** (Ann Arbor)

Redaktion: Paul **Lüttinger**

GESIS – Leibniz-Institut für Sozialwissenschaften
Postfach 12 21 55
68072 Mannheim
Tel.: 0621 – 1246-268
E-Mail: mda@gesis.org
Internet: www.gesis.org/MDA/

Die MDA deckt alle Fragestellungen aus dem Bereich der Empirischen Sozialforschung ab, insbesondere aus dem Bereich der Umfragemethodik. Im Vordergrund stehen Artikel, die die methodischen und/oder statistischen Kenntnisse der Profession erweitern, sowie Beiträge, die sich mit der Anwendung der Methoden der Empirischen Sozialforschung in der Forschungspraxis beschäftigen, oder solche, in denen ein statistisches Verfahren exemplarisch angewandt wird. Obwohl der Schwerpunkt auf Umfragemethoden liegt, sind Beiträge zu anderen methodischen Bereichen willkommen.

Alle Beiträge, die zur Veröffentlichung in der MDA eingereicht werden, werden von mindestens zwei unabhängigen Gutachtern blind begutachtet.

Der Nachdruck von Beiträgen ist nach Absprache möglich. Die MDA erscheint zweimal im Jahr und steht als Printversion und online zur Verfügung. Die Registrierung für den Bezug der MDA erfolgt über die Web-Seiten von GESIS:

http://www.gesis.org/Publikationen/Zeitschriften/MDA/MDA_subscribe.asp

Druck: Concordia-Druckerei König oHG, Mannheim-Sandhofen
Gedruckt auf chlorfrei gebleichtem Papier.

ISSN 1864-6956

2. Jahrgang 2008 © GESIS, Mannheim, Dezember 2008

Inhalt

95 Editorial

FORSCHUNGSBERICHTE

97 States oder Traits? Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews?
Anna Schnauber und Gregor Daschmann

125 Der Einfluss von Kontrollüberzeugungen der Interviewer auf die Teilnahme an telefonischen Befragungen
Christian Weidmann, Patrick Schmich und Sabine Schiller-Born

149 Zufallsbefragung von Ausländern auf Basis des Ausländerzentralregisters. Erfahrungen bei der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“ (RAM)
Christian Babka von Gostomski und Monika Pupeter

179 On the Treatment of Non-Original Sample Members in the German Household Panel Study (SOEP)
Martin Kroh, Rainer Pischner, Matin Spieß und Gert G. Wagner

REZENSIONEN

199 Methoden kulturvergleichender Sozialforschung. Eine Einführung. Susanne Rippl und Christian Seipel, 2008
Jan van Deth

202 Fragebogen. Ein Arbeitsbuch. Rolf Porst, 2008
Jörg Blasius

204 Counterfactuals and Causal Inference. Methods and Principles for Social Research. Stephen L. Morgan und Christopher Winship, 2007
Markus Gangl

206 Ways of Knowing. Competing Methodologies in Social and Political Research. Jonathon W. Moses und Torbjørn L. Knutsen, 2007
Joachim Behnke

ANKÜNDIGUNGEN

- 210 Mobile Research Conference line up announced:
Dr. Liz Nelson to Keynote
- 211 European Labour Force Survey (EU-LSF) and
European Union Statistics on Income and Living
Conditions (EU-SILC): 1st European User Conference
- 214 Call for Papers: 6. Nutzerkonferenz „Forschung mit dem
Mikrozensus“ Analysen zur Sozialstruktur und
zum sozialen Wandel
-
- 215 AutorInnen, RezensentInnen, GutachterInnen 2008
- 216 Hinweise für unsere Autorinnen und Autoren

Editorial

Liebe Leserin, lieber Leser,

mit diesem Heft ist der zweite Jahrgang der MDA abgeschlossen. Im Namen der Herausgeber danke ich allen AutorInnen, RezensentInnen und GutachterInnen, die an diesem Jahrgang mitgewirkt haben. Wie jede Ausgabe der MDA, bietet dieses Heft einen vielfältigen Einblick in die zur Zeit im Bereich Umfrageforschung laufenden Forschungsarbeiten. Die ersten beiden Artikel beschäftigen sich mit dem Teilnahmeverhalten an Telefoninterviews; in den letzten beiden Beiträgen geht es um die Gewinnung bzw. Erweiterung von Stichproben.

Im ersten Artikel berichten Anna Schnauber und Gregor Daschmann von einer Studie, in der Teilnehmer und Verweigerer einer telefonischen Befragung nach den Gründen ihres Verhaltens befragt wurden. Nach Aussage der Studie spricht einiges dafür, dass vor allem zwei Faktoren über Teilnahme oder Verweigerung entscheiden: der Zeitpunkt des Kontakts und die Einstellung der Zielperson zu Umfragen. Da die Autoren nur sehr schwache Korrelationen zwischen dieser Einstellung und anderen soziodemographischen Variablen oder Persönlichkeitsmerkmalen finden, und der Zeitpunkt des Kontakts eher zufällig über Gruppen von Befragten streut, gehen die Autoren davon aus, dass Verweigerungen nicht zu systematischen Verzerrungen der Ergebnisse von (telefonischen) Befragungen führen.

Der zweite, von Christian Weidmann, Patrick Schmich und Sabine Schiller-Born verfasste Beitrag beschäftigt sich mit dem Einfluss des Interviewers, genauer gesagt mit den Kontrollüberzeugungen des Interviewers, auf die Teilnahmebereitschaft von Befragten in telefonischen Umfragen. Überprüft wird die Hypothese, dass Interviewer mit starken internen Kontrollüberzeugungen, also der Überzeugung, dass das eigene Handeln maßgeblich zum Erfolg oder zum Misserfolg beiträgt, erfolgreicher bei der Durchführung von Interviews sind, also eine höhere Responsequote erzeugen. Auf Basis der Daten des telefonischen Gesundheitssurveys des Robert Koch-Instituts von 2007 bestätigt sich die von den Autoren formulierte Hypothese.

In dem Beitrag von Christian Babka von Gostomski und Monika Pupeter werden die Möglichkeiten und Grenzen des Ausländerzentralregisters (AZR) als Auswahlrahmen für Studien über Migranten vorgestellt. Die Herausgeber haben lange diskutiert, ob dieser Beitrag in der MDA veröffentlicht

werden sollte. Nicht, weil die Qualität des Beitrags in Frage gestellt wurde; hier waren wir uns mit den Gutachtern einig, dass der Beitrag gut gemacht ist. Grund für die Diskussion war vielmehr die, auch von den Gutachtern benannte Problematik, dass das AZR nur dem Bundesamt für Migration und Flüchtlinge als Stichprobenrahmen zur Verfügung steht. Auch die Daten des Projekts stehen (zumindest zur Zeit) nicht uneingeschränkt für Reanalysen zur Verfügung; eigentlich eine Voraussetzung für eine Veröffentlichung in der MDA.¹ Nach intensiver Beratung haben die Herausgeber entschieden, den Artikel dennoch zu veröffentlichen, um so einerseits Informationen zur Beurteilung der Qualität der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“ zugänglich zu machen und andererseits einen weiteren Impuls für die Diskussion um mögliche Stichprobenrahmen für Umfragen unter Migranten zu geben.

Martin Kroh, Rainer Pischner, Martin Spieß und Gert G. Wagner dokumentieren in ihrem Beitrag eine Besonderheit des SOEP. Im Gegensatz zu anderen großen Haushaltspanelstudien verfolgt das SOEP alle Personen über 16 Jahre, die jemals in einem Zielhaushalt gelebt haben, einerlei wann oder warum die Personen die (ursprünglichen) Zielhaushalte verlassen haben. Die Autoren diskutieren die Gründe für diesen Stichprobenplan, sie beschreiben die angewandte Verfolgungsregel, berichten Teilnahmewahrscheinlichkeiten für Mitglieder der ursprünglichen Stichprobe und für neue Stichprobenelemente und nennen schließlich einige Fragestellungen, die sich auf Basis dieser einzigartigen Stichprobenkonstruktion mit den Daten des SOEP untersuchen lassen.

Zur Lektüre empfehlen möchte ich auch die Rezensionen dieses Hefts, die zwei deutschsprachige und zwei englischsprachige Monographien betreffen.

Ich hoffe, dass die Lektüre dieses Hefts der MDA für Sie informativ ist und Sie gegebenenfalls auch Anregungen für Ihre eigenen Forschungsarbeiten erhalten. Besonders freuen würde ich mich, wenn Sie uns Manuskripte dieser Forschungsarbeiten zur Veröffentlichung in der MDA anbieten würden. Ich wünsche Ihnen, auch im Namen der Herausgeber, ein erfolgreiches neues Jahr.

Mannheim, 15. Dezember 2008

CHRISTOF WOLF

1 Die Autoren der Studie sind mit diesen kritischen Punkten von Beginn an offen umgegangen und haben sie bereits bei der ersten Einreichung des Artikels angesprochen.

States oder Traits?

Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews?

States or Traits?

Factors Influencing the Willingness to Participate in Telephone Surveys

Anna Schnauber und Gregor Daschmann

Zusammenfassung

Um zu untersuchen, was die Teilnahmebereitschaft an telefonischen Interviews beeinflusst, wurden 301 Teilnehmer und 326 Verweigerer einer telefonisch durchgeführten Marktforschungsstudie befragt. Basierend auf den geführten Interviews zeigt sich, dass in erster Linie zwei Faktoren entscheidend für die Teilnahmebereitschaft an telefonischen Umfragen sind: die persönliche Einstellung zu Umfragen und der Zeitpunkt der Befragung. Die Bedeutung der Zeit für die Teilnahmeentscheidung spricht für einen zufälligen Ausfallprozess – wird ein potenziell Befragter zu einem anderen Zeitpunkt kontaktiert, fällt seine Entscheidung wahrscheinlich auch anders aus. Die Umfrageeinstellung hingegen ist ein stabiler Einflussfaktor. Da sich aber nur wenige und schwache Zusammenhänge mit grundlegenden Persönlichkeitseigenschaften und soziodemografischen Merkmalen zeigen, spricht dies zwar dafür, dass es bestimmte Personen gibt, die Befragungen gegenüber grundsätzlich abgeneigt sind, diese sich aber nicht grundlegend von den Teilnehmern einer Befragung unterscheiden. Somit kann davon ausgegangen werden, dass Verweigerungen nicht zu systematischen Verzerrungen der Ergebnisse von Umfragen führen.

Abstract

To examine the factors influencing the willingness of target persons to participate in telephone surveys, 301 participants and 326 non-respondents of a market research telephone survey were interviewed on this subject. These interviews show that two factors are predominantly decisive for the willingness to participate in telephone surveys: the personal attitude towards surveys and the moment of interview. The significance of the moment for the decision on participation speaks for a random dropout process – if a potential interviewee is contacted at another time, he will probably make a different decision. By contrast, the attitude towards surveys is a stable influencing factor. However, there are only a few and weak correlations with primary personal characteristics and socio-demographic criteria; this speaks for there being certain persons who, in principle, are averse to surveys, but they do not differ essentially from the participants of a survey. Therefore it can be assumed that non-response does not automatically lead to systematic distortions of survey results.

1 Einleitung¹

Die Ausschöpfungsquoten bei Befragungen – vor allem bei telefonischen – sinken immer weiter (vgl. z. B. De Leeuw/De Heer 2002; Van der Vaart u. a. 2005); für die Sozialwissenschaft sind repräsentative Daten aus Umfragen jedoch unverzichtbar, denn sie stellen meist den einzigen Weg dar, verallgemeinerbare Informationen über Menschen, ihre Einstellungen, Meinungen und Handlungen zu generieren. Die sinkenden Quoten sind somit alarmierend und vor allem auf die steigende Zahl von Verweigerern zurückzuführen (vgl. z. B. Meier u. a. 2005; Hansen 2006). Deshalb ist die Gruppe der Verweigerer – ihre Merkmale, Gründe und Entscheidungsprozesse – besonders interessant und bedarf großer Aufmerksamkeit. Sollte sie sich grundsätzlich von den Teilnehmern an Befragungen unterscheiden, sind Verzerrungen in den Umfrageergebnissen wahrscheinlich. Die vorliegende Nonresponse-Studie untersucht daher die Merkmale und Motive von Verweigerern. Hierzu wurden im Rahmen einer repräsentativen Bevölkerungsumfrage 301 Befragungsteilnehmer sowie 326 Verweigerer (letztere mittels einer gesonderten Nachfassaktion) zu den Gründen ihrer Teilnahme bzw. Verweigerung befragt.

2 Hintergrund

Zum Thema Nonresponse und speziell zu den Mechanismen und Faktoren, die zur Verweigerung führen, wird bereits seit den 1940er-Jahren umfassend publiziert (vgl. Smith 2002). Die Forschung verläuft allerdings zum großen Teil unsystematisch und nicht theoriegeleitet. Um die bisher vorliegenden Ergebnisse zu möglichen Einflussfaktoren zu systematisieren, werden sie im Folgenden in zwei Gruppen eingeteilt: ‚Traits‘ und ‚States‘. Als Traits werden stabile Einstellungen und Verhaltensweisen bezeichnet, die unabhängig von einer konkreten Situation bzw. Interviewanfrage vorhanden sind. Unter States werden situative Faktoren, die im Moment der Kontaktaufnahme auftreten, verstanden.

Bezüglich der Traits zeigen sich in vielen Studien Zusammenhänge zwischen soziodemografischen Merkmalen der potenziell Befragten wie Alter, Geschlecht und sozioökonomischem Status und der Teilnahmewahrscheinlichkeit (vgl. z. B. Erbslöh/Koch 1988; Reuband/Blasius 2000), auch wenn diese nicht als ursächlich

1 Die Autoren danken der forum! Marktforschung GmbH in Mainz, insbesondere deren Inhaber und Geschäftsführer Roman Becker, für die Kooperation und vielfältige Unterstützung, ohne die die empirische Umsetzung der Studie nicht möglich gewesen wäre.

betrachtet werden können, sondern lediglich als ‚Globalvariablen‘ (Schnell 1997: 199) auf hinter ihnen liegende Mechanismen hinweisen. Diese implizit angenommenen Mechanismen im Hintergrund sind kaum empirisch untersucht. Dazu gehören z. B. grundlegende Persönlichkeitseigenschaften und Verhaltensweisen wie soziale Isolation, generelles Misstrauen oder Introvertiertheit. Vermutet und teilweise bestätigt wird, dass diese einen direkten Einfluss auf die Teilnahmeentscheidung haben (vgl. z. B. Esser 1973; Goyder 1987). Darüber hinaus liegt nahe, dass auch die Einstellung zu und die Erfahrung mit Umfragen die Entscheidung über eine mögliche Teilnahme beeinflussen. Vermutlich werden schlechte Erfahrungen mit Befragungen in der Vergangenheit zu größerem Misstrauen und damit zu einer niedrigen Teilnahmebereitschaft des potenziell Befragten führen (vgl. DeMaio 1980: 231; Schräpler 2000: 118; Groves/Lyberg 2001: 204) – so z. B. mit als Interviews getarnten Verkaufsanrufen bzw. –besuchen. Bereits seit den 1960er-Jahren ist dieses Problem bekannt; in einer Studie in den USA gaben 60% der Befragten an, mindestens einmal einen als Interview getarnten Verkaufsanruf erhalten zu haben (vgl. Biel 1967: 298). In einer jährlich durchgeführten Umfrage zur Akzeptanz von Interviews in Deutschland gaben 2006 fast zwei Drittel der Befragten an, bereits von einer Telemarketingagentur angerufen worden zu sein; 40% dieser Anrufe waren als Interviews getarnt (vgl. forsa 2006: 3). Dass dies zu sinkender Teilnahmebereitschaft bei folgenden Kontaktaufnahmen führt, zeigt Sheets und Kollegen in einem Experiment (Sheets u. a. 1974: 262).

Des Weiteren zeigen einige Studien einen Einfluss der Befragungshäufigkeit auf die Teilnahmebereitschaft: Sogenanntes ‚Oversurveying‘, die Überladung mit Umfrageanfragen, führt zu sinkender Beteiligung (vgl. Goyder 1987: 153; Groves/Cialdini/Couper 1992: 477; Groves/Couper 1998: 31).

Neben konkreten eigenen Erfahrungen mit Umfragen kann auch das Image, das jemand von Umfragen hat, einen Einfluss auf die Teilnahmebereitschaft haben. So führt eine positive Einstellung zu Umfragen zu einer erhöhten Teilnahmebereitschaft (vgl. Stocké/Becker 2004: 95). Umfragen über Umfragen zeigen in diesem Bereich, dass ihre Teilnehmer aus eigenem Interesse und Spaß mitmachen (vgl. Schleifer 1986: 24; Stocké/Langfeldt 2003: 7). Ein Negativfaktor bei der Beurteilung des Images von Befragungen ist vor allem mangelnder Glaube an die Anonymität von Umfragen (vgl. Esser 1973: 96; Esser 1986: 41) und die Sorge vor einem Eindringen in die Privatsphäre (vgl. Hartmann/Isaacson/Jurgell 1968: 297; Esser 1973: 96; DeMaio 1980: 231; Smith 1984: 475; Goyder 1986: 32; ähnlich Schleifer 1978).

Unabhängig von persönlichen Interessen kann auch die Überzeugung von der Wichtigkeit und Nützlichkeit von Umfragen eine Rolle spielen (vgl. Erbslöh/Koch 1988: 42). So beteiligen sich Menschen aus altruistischen Motiven an Um-

fragen, weil sie davon überzeugt sind, dass sie damit z. B. der Gesellschaft, der Wissenschaft oder allgemein demjenigen, der die Umfrage durchführt, helfen und damit beispielsweise Verbesserungen herbeiführen oder Ähnliches (vgl. Esser 1973: 94; Schleifer 1986: 24; Porst/von Briel 1995: 10).

Neben diesen stabilen Persönlichkeitsmerkmalen und Dispositionen untersuchen viele Forscher den Einfluss konkreter Befragungsmerkmale wie Länge und Thema auf die Teilnahmebereitschaft, meist mit dem Ergebnis, dass ein enger Zusammenhang zur Teilnahmebereitschaft besteht (vgl. z. B. Groves/Presser/Dipko 2004; Hansen 2006). Erhebungsart und durchführende Institution bzw. Auftraggeber der Studie können ebenfalls als Einflussgrößen wirken (vgl. z. B. Groves/Lyberg 2001; Tuckel/O'Neill 2002).

Die Entscheidung über die Teilnahme wird beim telefonischen Interview innerhalb weniger Momente getroffen. Deshalb ist es unwahrscheinlich, dass sie (nur) auf einer systematischen Verarbeitung und Abwägung der genannten Einflussfaktoren beruht. Es ist wahrscheinlicher, dass sie als Low-Involvement-Situation stark heuristisch beeinflusst ist. Dementsprechend ist anzunehmen, dass auch Reize, die unmittelbar in der Anfragesituation wirken, eine Rolle spielen (vgl. Schnell 1997; Groves/Couper 1998). Neben der in der Fachliteratur kaum untersuchten Situation des Befragten, geprägt durch Stress, Zeitdruck und Stimmungslage (vgl. z. B. Esser 1986; Goyder 1987), wirkt sich hier auch die Person des Interviewers aus. Sowohl seine soziodemografischen Merkmale wie Geschlecht und Alter als auch seine Erfahrung als Interviewer können Einfluss ausüben. Die Selbsteinschätzung bzw. Erwartungshaltung der Interviewer hat ebenfalls einen Einfluss auf die Kooperationsrate. Geht ein Interviewer selbstbewusst in das Interview und ist davon überzeugt, die potenziell Befragten zur Teilnahme motivieren zu können, erhält er weniger Verweigerungen (vgl. Singer/Frankel/Glassman 1983: 74; Schnell 1997: 196f.; Groves/Couper 1998: 36). Über seine Stimme und Sprache vermittelt er dem potenziell Befragten ein Bild von sich und seinem Anliegen, das ebenfalls ein Einflussfaktor sein kann (vgl. z. B. Singer/Frankel/Glassman 1983; Oksenberg/Cannell 2001).

Es ergibt sich somit aus dem Forschungsstand, dass eine Studie zu den Ursachen der Teilnahmebereitschaft in telefonischen Interviews folgende Faktoren berücksichtigen muss (siehe Abbildung 1 für Details):

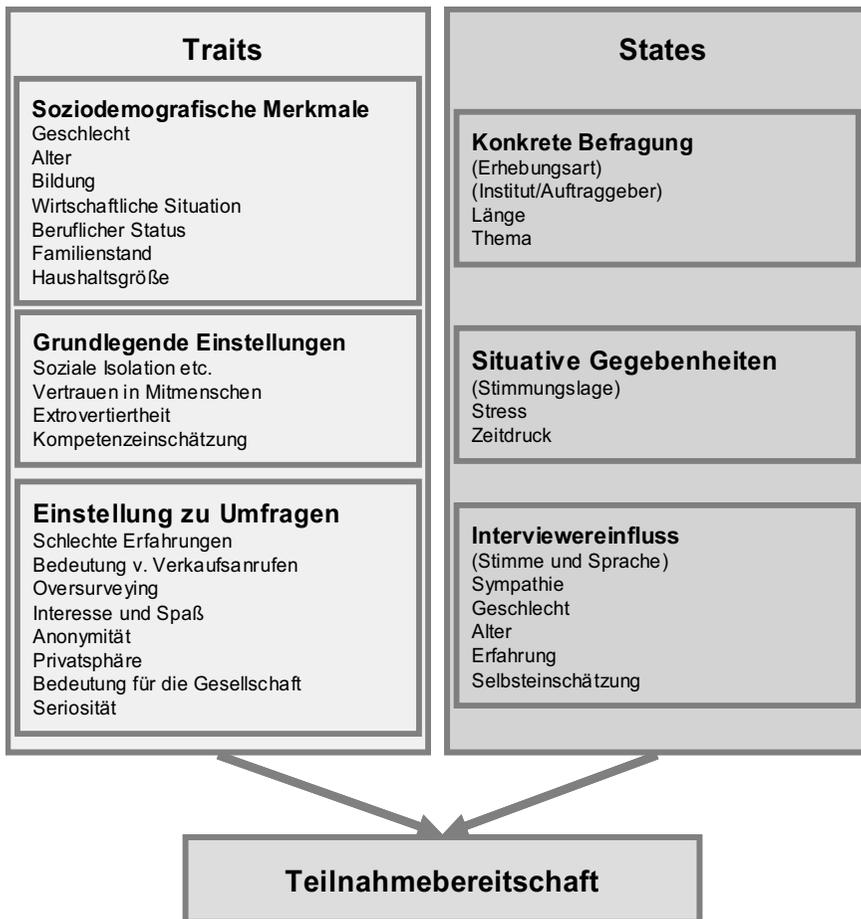
Traits:

- soziodemografische Merkmale,
- grundlegende Persönlichkeitseigenschaften und Verhaltensweisen,
- Einstellungen zu und Erfahrungen mit Umfragen.

States:

- Merkmale der konkreten Befragung,
- situative Gegebenheiten wie Zeitdruck und Stress des potentiellen Befragungsteilnehmers,
- Interviewermerkmale.

Abbildung 1 Einflussfaktoren auf die Teilnahmeentscheidung an Befragungen²



2 Die in Klammern stehenden möglichen Einflussfaktoren konnten in der vorliegenden Studie nicht empirisch untersucht werden, vgl. die Ausführungen zur Methode.

Grundsätzlich sind alle diese Einflussfaktoren nicht unproblematisch: Sie implizieren, dass das Teilnahmeverhalten innerhalb der zu befragenden Population nicht zufällig streut, sondern systematisch an andere Merkmale gekoppelt ist. Das bedeutet in der Konsequenz: Befragte unterscheiden sich von Nicht-Befragten nicht nur hinsichtlich der Teilnahmebereitschaft, sondern auch in anderen Eigenschaften und Merkmalen. Daraus wäre zu folgern, dass sämtliche Befragungstichproben in ihrer Zusammensetzung systematisch verzerrt sind und somit auch mit ihren Ergebnissen ein über den üblichen Stichprobenfehler hinausgehendes verzerrtes Bild der Grundgesamtheit zeichnen – zumindest bezüglich dieser Erklärungsvariablen. Die Relevanz dieser inhaltlichen Verzerrung hängt somit davon ab, wie zufällig oder systematisch diese Erklärungsvariablen mit den zentralen Erkenntnisparametern der jeweiligen Befragung zusammenhängen. Je nach Konstellation hat dies unterschiedliche Konsequenzen für die Forschung:

Sind vor allem grundlegende Einstellungen und Charaktereigenschaften – indirekt auch soziodemografische Merkmale, die auf hinter ihnen liegende Mechanismen hinweisen – entscheidungsbeeinflussend, stellt dies die Markt-, Meinungs- und Sozialforschung vor die größten Probleme. Denn hierauf haben die Forscher keinen Einfluss, es besteht kein Ansatzpunkt zur Steigerung der Teilnahmebereitschaft. Inhaltliche Verzerrungen sind dann wahrscheinlich, weil es sich um eine bestimmte Gruppe mit stabilen Persönlichkeitsmerkmalen handelt, die sich der Befragung grundsätzlich entzieht. Allerdings kann ein fundiertes Wissen über diese Gruppe dafür genutzt werden, Gewichtungungsverfahren zu entwickeln, die die inhaltlichen Verzerrungen nachträglich bestmöglich korrigieren.

Trägt die Erfahrung mit und die Einstellung zu Umfragen den größten Teil zur Entscheidung bei, bietet sich ein Ansatzpunkt für die Markt-, Meinungs- und Sozialforschungsbranche und auch für jedes einzelne Institut, gezielt durch eine Verbesserung des Images und damit der Vertrauens- und Glaubwürdigkeit, eine konsequente Abgrenzung von Direktmarketing, kontinuierliche Öffentlichkeitsarbeit etc. das Umfrageklima zu verbessern und so die Ausschöpfungsquoten zu erhöhen (vgl. Kaase 1999). Inhaltliche Verzerrungen sind hier allerdings kaum zu befürchten, da man davon ausgehen kann, dass negative Erfahrungen mit Kontaktaufnahmen bisheriger Befragungen zufällig über die Grundgesamtheit streuen, was aufgrund der meisten Stichprobenpläne anzunehmen ist. Sollten allerdings die Einstellungen zu Umfragen auf Persönlichkeitsmerkmalen des potenziell Befragten basieren, sind – wie bereits im vorausgegangenem Absatz ausgeführt – durchaus inhaltliche Verzerrungen möglich.

Haben Merkmale der konkreten Befragung wie die Erhebungsmethode, das Thema und die Länge den größten Einfluss, sind Verzerrungen der Ergebnisse wahr-

scheinlich. Denn wenn nur die am Thema Interessierten und diejenigen, die dafür auch bereit sind, viel Zeit aufzubringen, teilnehmen, findet eine Selbstselektion statt, die zu einem falschen Bild in den Ergebnissen führt. Hier kann der Forscher durch gezielte Designmaßnahmen wie die Kürzung des Interviews entgegenwirken. Für die Umfrageforschung allgemein sind hingegen keine gravierenden Konsequenzen zu befürchten.

Sind vor allem situative Faktoren wie der Zustand des potenziell Befragten zum Zeitpunkt der Interviewanfrage oder der Interviewer entscheidungsbeeinflussend, ist eine inhaltliche Verzerrung unwahrscheinlich, der Ausfallprozess ist stark zufallsgesteuert und damit nicht systematisch. Konvertierungsversuche zu einem anderen, für den potenziell Befragten passenden Zeitpunkt durch einen anderen Interviewer sind dann vielversprechend und können zusätzlich die Ausschöpfung erhöhen.

Zusammenfassend ist somit festzuhalten, dass sich erst durch die Identifikation der Haupteinflussgrößen auf die Teilnahmebereitschaft entscheiden lässt, ob durch Verweigerungen mit starken inhaltlichen Verzerrungen zu rechnen ist und welche Maßnahmen zum Umgang mit Nonresponse die größten Verbesserungen versprechen. Abbildung 1 fasst die bisherigen Überlegungen noch einmal systematisch zusammen.

3 Modell der Hierarchie der Einflussgrößen auf die Teilnahmebereitschaft

Neben der Identifikation verschiedener Einflussgrößen ist ein Modell des Prozesses, der im Moment der Interviewanfrage abläuft, zur Strukturierung hilfreich. Dabei soll das Modell nicht den Anspruch der Allgemeingültigkeit erheben, sondern als idealtypischer Modellfall bzw. als Taxonomie zum besseren Verständnis des Prozesses dienen, indem es eine denkbare Hierarchie der Einflussgrößen auf die Teilnahmebereitschaft verdeutlicht. Diese Hierarchie dient zur Gliederung der Ergebnisdarstellung; ein empirischer Test der vom Modell implizierten Effekthierarchien ist jedoch im Rahmen der vorliegenden Studie nicht geplant.

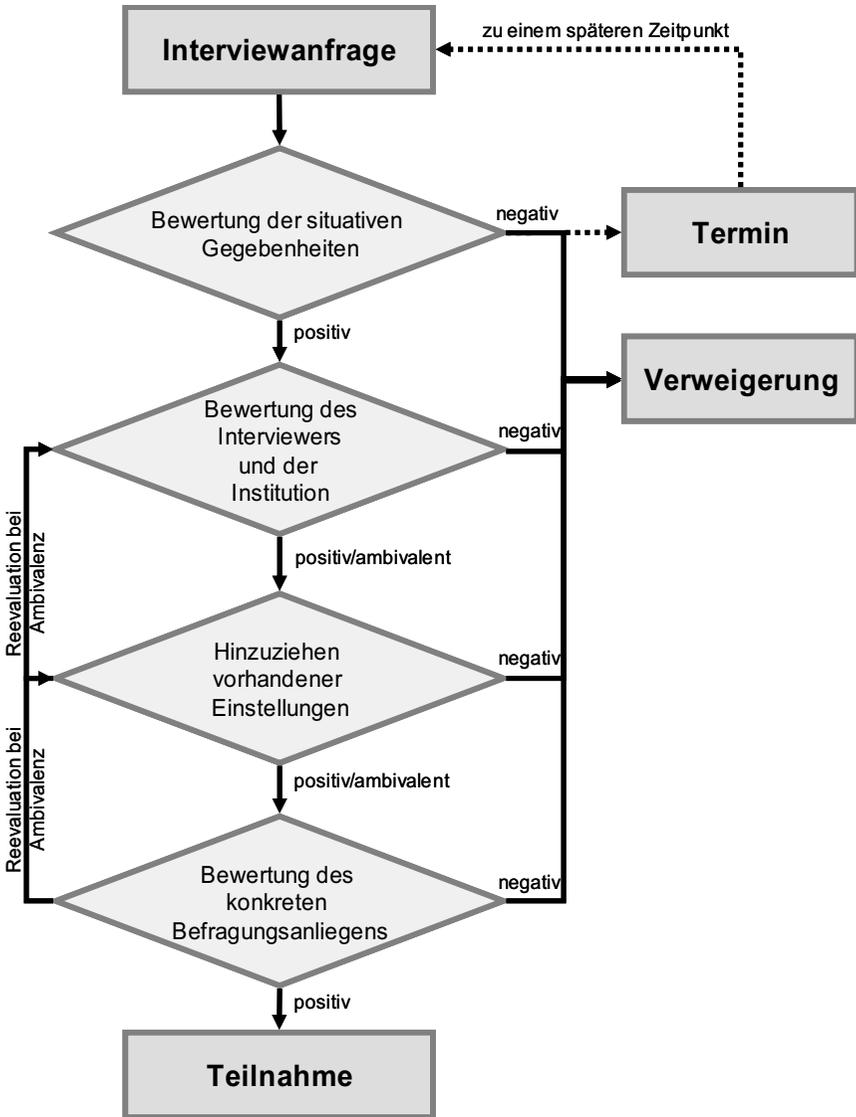
Hierzu gibt es in der Nonresponse-Literatur einige Ansätze; die meisten davon sind allerdings kaum theoriegeleitet. Ausnahmen stellen vor allem Schnell (1997) und Groves und Kollegen (verschiedene Veröffentlichungen) dar. Angelehnt an Schnell und Groves/Couper (1998) kann das Teilnahmeverhalten als rationale Entscheidungsfindung in einer Low-Cost-Situation betrachtet werden (vgl. Schnell 1997: 198; Groves/Couper 1998: 32). Demnach findet im Moment der Interviewanfrage eine Kosten-Nutzen-Abwägung statt, der potenziell Befragte sucht also nach

der Handlungsalternative – Teilnahme oder Verweigerung –, die ihm in der konkreten Situation den größten Profit bzw. die ‚stärkste Bedürfnisbefriedigung‘ (Schnell 1997: 159) verspricht. In diesen Prozess fließen sowohl Traits als auch States ein (vgl. Groves/Couper 1998). Da die Bitte zur Teilnahme an einer Befragung in den meisten Fällen eine Low-Involvement-Situation darstellt, der potenziell Befragte also kein Interesse daran hat, auf seine Entscheidungsfindung viele kognitive Ressourcen zu verwenden, läuft diese Kosten-Nutzen-Abwägung vermutlich nicht (komplett) bewusst, sondern heuristisch ab (vgl. Groves/Couper 1998). Außerdem wird der potenziell Befragte nicht alle Informationen auf einmal verarbeiten. Es gibt mehrere, aufeinander aufbauende Schritte. Jeder Schritt stellt eine eigene Entscheidungsfindung dar (vgl. Groves/Couper 1998). Fällt der erste Entscheidungsschritt negativ aus, verweigert der potenziell Befragte das Interview, fällt er positiv oder ambivalent aus, setzt sich der potenziell Befragte mit einem weiteren Schritt auseinander. Die Entscheidungsfindung in den verschiedenen Schritten kann ein unterschiedliches Maß an systematischer oder heuristischer Abwägung der Kosten und Nutzen erfordern.

Abbildung 2 stellt einen solchen idealtypischen Modellfall dar: Der *erste Schritt*, die Bewertung der situativen Gegebenheiten, stellt vermutlich eine Art Filter für die weiteren Schritte dar. Hier wird die Entscheidung sehr schnell und dementsprechend heuristisch ohne viel Mühe getroffen. Werden die situativen Gegebenheiten negativ bewertet, fühlt sich der potenziell Befragte also gestresst, steht er unter Zeitdruck und/oder ist er in schlechter Stimmung, wird er keine bzw. wenig Bereitschaft zeigen, sich überhaupt mit dem Anliegen des Interviewers zu befassen.

Erst wenn der potenziell Befragte seine situativen Gegebenheiten positiv bewertet, kann eine Interaktion zwischen ihm und dem Interviewer stattfinden. Im *zweiten Schritt* bewertet der potenziell Befragte das Anliegen des Interviewers und – falls bekannt – die Institution, die hinter der Anfrage steht. Kommt der potenziell Befragte zu dem Ergebnis, dass es sich um eine ihm unbekannte oder unseriöse und wenig vertrauenswürdige Institution handelt oder wirkt der Interviewer unsympathisch und nicht vertrauenserweckend, verweigert er die Teilnahme. Die möglichen Kosten wie z. B. die Sorge vor einem Verkaufsanruf oder die missbräuchliche Verwendung seiner Daten sind zu vordergründig, der potenziell Befragte wird keine Bereitschaft zeigen, einen möglichen Nutzen, der in dieser Situation weniger offensichtlich ist, in seine Abwägung mit einzubeziehen.

Abbildung 2 Modell der Hierarchie der Einflussgrößen auf die Teilnahmebereitschaft



Im *dritten Schritt* zieht er seine grundlegenden Einstellungen und Verhaltensweisen hinzu. Dies erfordert eine intensivere Beschäftigung mit dem Anliegen des Interviewers und ist dementsprechend erst zweckmäßig, wenn der potenziell Befragte bereits Schritt 2 positiv bewertet hat. Hier spielen einmal Persönlichkeitsmerkmale wie soziale Isolation oder Kommunikationsfreude eine Rolle, aber auch Einstellungen zu und Erfahrungen mit Umfragen. Im *vierten Schritt* wägt der potenziell Befragte ab, welche Kosten und welchen Nutzen die konkrete Befragung für ihn haben kann. Einflussfaktoren sind hier z. B. Länge und Thema.

Die Schritte zwei, drei und vier sind durch eine Art ‚Rückkopplung‘ miteinander verbunden: Wird einer der Schritte ambivalent bewertet, so wägt der potenziell Befragte zusätzlich den darauffolgenden Schritt ab. Aufbauend auf dieser Abwägung reevaluiert er den vorherigen Schritt. Ist er z. B. zunächst unschlüssig, ob er dem Interviewer und/oder der hinter ihm stehenden Institution vertrauen kann, zieht er seine vorhandenen Einstellungen und Erfahrung heran. Hat er bereits schlechte Erfahrungen gemacht, wird er auf dieser Basis erneut abwägen, wie er mit der Anfrage umgeht, und Schritt zwei negativ bewerten.

Im Gegensatz zu den anderen in Abbildung 1 dargestellten Einflussfaktoren finden sich die soziodemografischen Merkmale als ‚Globalvariablen‘ (Schnell 1997: 199) in diesem Modell der Teilnahmebereitschaft nicht direkt wieder. Niemand verweigert z. B., weil er alt, eine Frau oder ungebildet ist. Dementsprechend können die Soziodemografika als ‚Grundstock‘ gesehen werden: Nur die Persönlichkeitsmerkmale, Verhaltensmuster und Lebenssituationen, die mit ihnen zusammenhängen, üben einen direkten Einfluss aus. Diese finden sich in allen Schritten des Modells wieder.

4 Methode

4.1 Erhebungsmethode

Die Untersuchung von Einflussfaktoren auf die Teilnahmebereitschaft stellt eine besondere Herausforderung dar, weil die Nonrespondenten gerade dadurch definiert sind, dass über sie keine (näheren) Informationen vorliegen und besonders die Verweigerer auch nicht bereit sind, diese preiszugeben. Zur Identifikation von Einflussfaktoren auf die Teilnahmebereitschaft ist es aber wichtig, über die gleichen Informationen sowohl von den Teilnehmern als auch von den Verweigerern zu verfügen, denn nur über den Gruppenvergleich können Unterschiede identifiziert und ihre Einflusstärke bestimmt werden. Dies kann nur eine genuine Nonresponse-Studie leisten. Durch eine Nachbefragung wird gezielt versucht, Informationen

über Eigenschaften, Einstellungen oder Verhaltensweisen der Ausfälle zu sammeln, von denen man annimmt, dass sie mit ihrer (fehlenden) Teilnahmebereitschaft zusammenhängen (vgl. Smith 1983). Hauptschwachpunkt dieser Methode ist, dass auch Nonresponse-Studien immer mit einem nicht unerheblichen Nonresponse behaftet sind.³ Deshalb ist zur Interpretation der Ergebnisse die Annahme nötig, dass diejenigen, die sich bereit erklären, an der Verweigererbefragung teilzunehmen, eher den endgültigen Verweigerern und nicht den Teilnehmern ähnlich sind (vgl. Erbslöh/Koch 1988). Ergebnisse aus Nonresponse-Studien sind also mit einem Unsicherheitsfaktor belastet, denn über diejenigen, die auch die Nachbefragung verweigern bzw. nicht erreichbar sind, liegen anschließend immer noch keine oder nur wenige Informationen vor. Deshalb sind Ergebnisse aus Nonresponse-Studien als Tendenzen zu verstehen.

Die Hauptbefragung, deren Teilnehmer und Verweigerer im Folgenden untersucht werden, ist eine telefonisch durchgeführte Marktforschungsstudie.⁴ Die Stichprobe ist eine nach Bundesländern geschichtete zufallsgenerierte Telefonstichprobe.⁵ Ein Teil der Stichprobe besteht aus Panelkontakten, also Personen, die im vergangenen Jahr bereits teilgenommen und einer erneuten Befragung zugestimmt haben. Bei diesen Kontakten war zunächst kritisch zu prüfen, ob sie in die folgenden Analysen einbezogen werden können, denn es war zu vermuten, dass sie sich in ihrem Teilnahme- und Verweigerungsverhalten systematisch von den erstmals Kontaktierten unterscheiden. Da sich jedoch insgesamt zeigte, dass es wenige Unterschiede zwischen den Panel- und den Erstkontakten gab, wurden diese in die Analyse mit einbezogen und alle auftretenden Unterschiede zwischen den beiden Gruppen an entsprechender Stelle erwähnt.

Die vorliegende Studie berücksichtigt zwei Kritikpunkte an der bisherigen Methodenliteratur: Erstens werden die Einflussfaktoren anhand einer kommerziellen Marktforschungsstudie untersucht und nicht anhand von akademischen oder staatlichen Umfragen, denn diese haben insgesamt weniger mit Verweigerungen

- 3 Erbslöh/Koch berichten für die Nonresponse-Studie zum ALLBUS 1986 von einer Ausschöpfungsquote von 44% (vgl. Erbslöh/Koch 1988), Smith berichtet von 53% in einer vom Bureau of Social Science Research 1981 durchgeführten Verweigererbefragung und von 29% in einer anderen Studie (vgl. Smith 1983).
- 4 Bei der zugrunde liegenden Hauptbefragung handelt es sich um eine Studie der forum! Marktforschung GmbH in Kooperation mit der Deutschen Gesellschaft für Qualität e. V., das sogenannte „ExBa (Excellence Barometer) Bevölkerung“. Das Excellence Barometer wird seit 2001 jährlich durchgeführt. Befragt wurden 1.000 Personen über 18 Jahre. Die angekündigte Interviewlänge betrug 12 bis 15 Minuten, Thema waren „Fragen zu verschiedenen Themen des täglichen Lebens“.
- 5 Zum Verfahren zur Ziehung von zufallsgenerierten Telefonstichproben in Deutschland vgl. Häder/Gabler (1998) und Rösch (1998).

zu kämpfen (vgl. Esser 1986: 43; Stocké/Becker 2004: 93f.). Zweitens werden die Mechanismen einer Telefonbefragung untersucht. Telefonische Interviews sind heute die meisteingesetzte Methode in der kommerziellen Forschung (vgl. Brosius/Koschel 2001; Diekmann 2006).

4.2 Datenquellen

Um möglichst viele der Einflussfaktoren erfassen zu können, ist der Einsatz verschiedener Quellen erforderlich:

Erstens werden Informationen über die Interviewer in die Analyse einbezogen. Geschlecht, Alter und Interviewerfahrung (gemessen an der Dauer der Arbeit als Interviewer in Jahren) sind bereits von jedem Interviewer bekannt und können – da die Studie CATI-gestützt durchgeführt wird – mit den Befragungsdaten verknüpft werden. Um die Selbsteinschätzung der Interviewer mit einzubeziehen, werden alle Interviewer im Anschluss an die Schulung der Hauptbefragung darum gebeten, auf einer Skala von 1 ‚sehr schwierig‘ bis 5 ‚überhaupt nicht schwierig‘ anzugeben, für wie schwierig sie es halten, einen Ansprechpartner zur Teilnahme an der Hauptbefragung zu motivieren.

Die wichtigen Eigenschaften Stimme und Sprache können leider nicht in die vorliegende Untersuchung einbezogen werden. Aufgrund der Größe der Stichprobe und der fehlenden technischen Mittel zu Aufzeichnung aller Kontaktversuche ist es nicht möglich, jede Interaktion nach linguistischen Kriterien zu bewerten. Damit fehlt in Bezug auf den Interviewereinfluss eine entscheidende Determinante, denn Stimme und Sprache sind das einzige Medium, das zur Verfügung steht, um die Verbindung zwischen potenziell Befragtem und Interviewer sowie Befragung herzustellen. Nur über seine Stimme und Sprache vermittelt er dem potenziell Befragten ein Bild von sich und seinem Anliegen.

Um zumindest zu erfassen, wie der Interviewer auf den Befragten nach dessen subjektiver Einschätzung gewirkt hat, werden die Befragten im Fragenblock (siehe unten) gebeten anzugeben, wie entscheidend das Verhalten des Interviewers für ihre Entscheidung, teilzunehmen oder zu verweigern, war: Sie bewerten anhand einer fünf-stufigen Zustimmungsskala die Aussage: „Entscheidend [für die Teilnahme/Verweigerung] war vor allem, wie sich der Interviewer verhalten hat („Faktor Interviewer“).“

Die zweite und wichtigste Quelle stellt ein Fragenblock dar, den sowohl Teilnehmer als auch Verweigerer beantworten: er erfasst die in der Literatur als relevant identifizierten möglichen Faktoren. Dieser Fragebogen muss mehreren Anforderungen gerecht werden: Erstens darf er aufgrund des Studiendesigns nicht lang sein. Die Wahrscheinlichkeit, dass ein Verweigerer, also eine nicht befragungswillige

Person, sich dennoch bereit erklärt, viel Zeit zu investieren, ist gering. Darüber hinaus ist es im Rahmen der Hauptbefragung – also für die Teilnehmerinterviews – nicht möglich, einen zusätzlichen langen Fragenblock zu integrieren. Zweitens darf der Fragebogen keine bzw. nicht viele Fragen enthalten, die als zu persönlich empfunden werden können. Wenn den Verweigerern angekündigt wird, dass sie zu ihrer Verweigerungsentscheidung befragt werden sollen, müssen sich die Fragen auch hauptsächlich darauf beziehen, sonst sind Abbrüche wahrscheinlich. Drittens sollte der Fragebogen möglichst viele der in der Literatur als relevant diskutierten Einflussfaktoren enthalten.

Der ca. vierminütige Fragebogen enthält eine Itembatterie mit zehn Aussagen, die die Bereiche grundlegende Einstellungen, Einstellungen zu Umfragen und Stress als situativen Faktor abdecken. Wo es möglich ist, werden Formulierungen aus ähnlichen Untersuchungen bzw. aus Umfragen über Umfragen übernommen, da bei ihnen davon auszugehen ist, dass sie verständlich sind:

- „Umfragen bringen Abwechslung und sind interessant“ (Stocké/Langfeldt 2003: 7) zur Messung des persönlichen Interesses und Spaßes an der Teilnahme an Befragungen;
- „Marktforschungsunternehmen behandeln die Daten vertraulich“ (vgl. Schleifer 1986: 24) zur Messung des Glaubens an die Vertraulichkeit der Datenbehandlung in der kommerziellen Forschung;
- „Bei Umfragen wird häufig etwas gefragt, was niemanden etwas angeht“ (Stocké/Langfeldt 2003: 7) zur Messung der Einstufung von Umfragen als Eingriff in die Privatsphäre;
- „Oft weiß man gar nicht, ob mit einer Umfrage nicht eigentlich etwas verkauft werden soll“ zur Messung der Vermischung von Umfragen und Verkaufsanrufen;
- „Umfragen sind für die Gesellschaft wichtig und sinnvoll“ (vgl. Porst 1998: 13) zur Messung altruistischer Motive für die Teilnahme;
- „Bei den meisten Umfragen ist von vorneherein klar, was rauskommt“ (Porst 1998: 13) zur Messung des Vertrauens in die Glaubwürdigkeit, Ernsthaftigkeit und Seriosität von Umfragen;
- „Es gibt nur sehr wenige Menschen, auf die man sich verlassen kann“ zur Messung des Vertrauens in die Umwelt;
- „Ich sage meine Meinung immer und überall offen“ zur Messung der Extrovertiertheit des Befragten;
- „Ich spreche oft mit anderen Menschen über Themen des täglichen Lebens“ zur Messung der Kommunikativität und Kompetenzeinschätzung in Bezug auf die Interaktion;
- „Mein Tag war [heute/am Tag der Hauptbefragung] sehr stressig“ zur Messung des subjektiv empfundenen Zeitdrucks.

Des Weiteren enthält der Fragebogen Fragen zur Anzahl der Interviewanfragen und Teilnahme an Befragungen in den vergangenen zwölf Monaten. So zeigt sich, ob mit steigender Zahl von Umfrageanfragen die Teilnahmebereitschaft sinkt. Außerdem werden die Befragten gebeten, ihre letzte Umfragerfahrung als eher angenehm oder eher unangenehm einzustufen.

Schwierig ist die Messung des Einflusses von Thema und Länge der Hauptbefragung als Teilnahme- bzw. Verweigerungsgrund. Für eine verlässliche Messung müssten diese beiden Faktoren experimentell variiert werden; da dies im Rahmen dieser Untersuchung nicht möglich ist,⁶ sollen die Befragten anhand einer fünfstufigen Zustimmungsskala bewerten, wie wichtig diese Faktoren für ihre Entscheidung waren. Dies kann zumindest als Indiz für den Einfluss dieser beiden Faktoren gesehen werden. Zur besseren Verständlichkeit für die beiden Gruppen werden die Aussagen angepasst formuliert; so lautet die Aussage im Verweigerer-Fragenblock: „Entscheidend war vor allem, dass das Thema uninteressant war“, im Teilnehmer-Fragenblock hingegen: „Entscheidend war vor allem, dass das Thema interessant war“. Die Länge wird folgendermaßen erfasst: „Entscheidend war vor allem, wie lange das Interview dauern sollte.“

Um neben dem situativen Faktor ‚Stress‘, der im Rahmen der Itematterie abgefragt wird, auch die zeitliche Komponente zu erfassen, bewerten die Befragten anhand einer fünfstufigen Zustimmungsskala, ob es entscheidend für ihre Teilnahme- bzw. Verweigerungsentscheidung war, dass sie gerade Zeit hatten (Teilnehmer) bzw. dass sie gerade keine Zeit hatten (Verweigerer).

4.3 Studiendesign

Sowohl die Teilnehmer⁷ als auch die Verweigerer der Hauptbefragung sollen den Fragenblock beantworten. Während dies bei den Teilnehmern problemlos möglich ist, indem der Fragenblock in die Hauptbefragung⁸ integriert wird, besteht die Schwierigkeit darin, auch die Verweigerer zur Beantwortung des Fragenblocks zu motivieren.

Um eine möglichst große Anzahl an Verweigerern zur Beantwortung des Fragenblocks zu bewegen, wurde folgende mehrstufige Vorgehensweise gewählt:

6 Die Nonresponse-Studie ist an eine reguläre kommerzielle Marktforschungsstudie angegliedert, deren Thema und Länge nicht variabel sind; vgl. Fußnote 4.

7 Um einen unverhältnismäßig hohen Aufwand zu vermeiden, werden nicht alle Teilnehmer befragt, sondern lediglich eine Zufallsstichprobe von ca. 300 Teilnehmern.

8 Die Fragen werden nach dem inhaltlichen Teil der Hauptbefragung, aber vor den soziodemografischen Fragen gestellt.

Verweigert ein potenziell Befragter die Teilnahme an der Hauptbefragung, versucht der Interviewer ihn – z. B. mit Hinweis auf die Kooperation mit der Johannes Gutenberg-Universität Mainz und die verkürzte Befragungsdauer – für die Beantwortung des o. g. Fragenblocks zu gewinnen.

Um ein möglichst vollständiges Bild von der Gesamtgruppe der Verweigerer zu erhalten, wird ein Teil der Personen, die auch die Zusatzbefragung verweigert haben bzw. bei denen der Interviewer keine Möglichkeit zur Nachfrage hat, nochmals telefonisch kontaktiert und unter Hinweis auf das Ziel, Aufschlüsse über Verweigerungsverhalten zu erhalten, befragt.

4.4 Durchführung

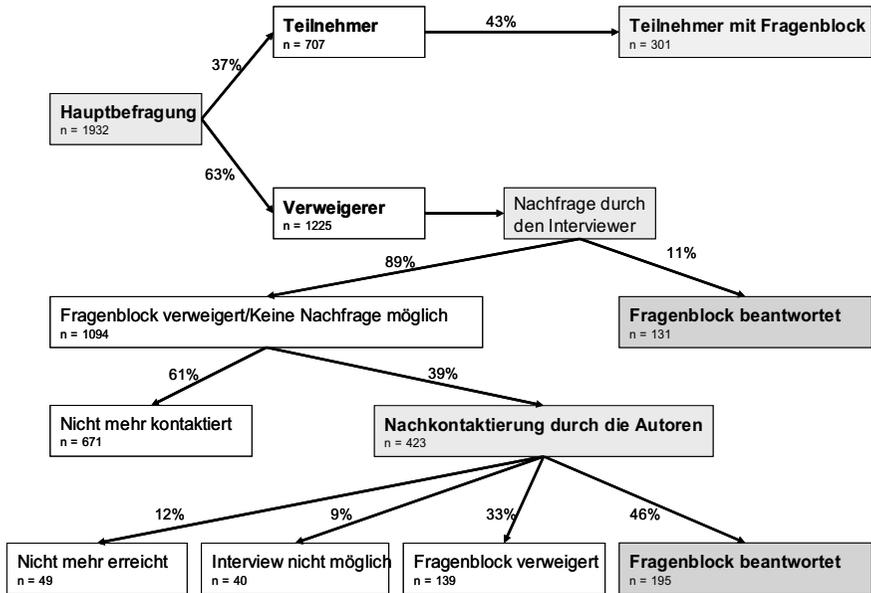
Die Interviews – sowohl mit den Teilnehmern als auch mit den Verweigerern – wurden im Zeitraum vom 12. bis 30. März 2007 werktags zwischen 17 und 21 Uhr aus dem Interview-Center der forum! Marktforschung GmbH in Mainz von 18 Interviewern und den Autoren geführt. In einer Schulung wurde den Interviewern das Vorgehen erläutert. Abbildung 3 stellt den Ablauf der Studie und die Fallzahlen in den einzelnen Gruppen dar.

Insgesamt beantworteten 326 Verweigerer den Fragenblock. Das entspricht 27% aller Verweigerer im Rahmen der Hauptbefragung. Betrachtet man weiterhin die 671 nicht erneut durch die Autoren kontaktierten Verweigerer als neutralen ‚Ausfall‘, da sie nicht Teil der Zufallsstichprobe waren, wurde mit 59% aller (mehrfach kontaktierten) Verweigerer der Hauptbefragung ein Verweigererinterview geführt. Vergleichbare Studien in der Vergangenheit erreichten mit 29 bis 53% niedrigere Quoten (vgl. Smith 1983; Erbslöh/Koch 1988).

139 Verweigerer verweigerten explizit auch auf Nachfrage die Beantwortung des Fragenblocks; damit birgt auch diese Nonresponse-Studie ein Verzerrungspotenzial und ihre Ergebnisse können nur als Tendenzen verstanden werden. Während über die 139 ‚Komplettverweigerer‘ als Ganzes kaum Informationen vorliegen, sind von einem kleinen Teil dieser Gruppe (24 Personen) detaillierte soziodemografische Informationen bekannt,⁹ die aber nur wenige, schwache Zusammenhänge mit der Teilnahmebereitschaft aufweisen. Um die entscheidenden Einflussfaktoren zu identifizieren, ist eine reine Gegenüberstellung anhand von soziodemografischen Merkmalen nicht ausreichend. Erst die Auswertung des Fragebogens macht dies möglich. Damit stellen die 301 Teilnehmer- und 326 Verweigererinterviews die Basis für die folgenden Ergebnisse dar.

9 Hierbei handelt es sich um Panelkontakte, die im vergangenen Jahr an der zugrunde liegenden kommerziellen Marktforschungsstudie teilgenommen hatten.

Abbildung 3 Studiendesign mit erreichten Fallzahlen



5 Ergebnisse

Die in der Nonresponse-Forschung am häufigsten untersuchten Zusammenhänge mit soziodemografischen Merkmalen lassen sich nur zum Teil bestätigen und erweisen sich durchgängig als sehr schwach. Den stabilsten Zusammenhang zeigt das Geschlecht: Frauen verweigern häufiger als Männer ($\chi^2 = 13,0$, Cramer-V = 0,144, $p < 0,001$). Tendenziell ist auch ein Alterseffekt erkennbar – Alte verweigern eher als Junge –, dieser ist aber nicht unter allen Bedingungen messbar und nicht signifikant. Da dieser Effekt sich aber am stärksten (wenn auch nicht signifikant) unter den Erstkontakten zeigt, spricht dies für den erwarteten Alterseffekt, dass alte Menschen eher verweigern als Jüngere. Auch der sogenannte Mittelschicht-Bias,¹⁰ der in der Literatur oft diskutiert wird, kann durch die vorliegenden Daten nicht bestätigt werden: Bildung und wirtschaftliche Lage haben keinen (stabilen) Einfluss auf die Teilnahmebereitschaft. Dies gilt gleichermaßen für Panel- und Erstkontakte.

10 Nach dieser These entziehen sich besonders die unteren sozialen Schichten – also die niedrig Gebildeten und wirtschaftlich schlechter Gestellten – der Teilnahme an Befragungen (vgl. z. B. Esser 1973; Erbslöh/Koch 1988).

Teilnehmer und Verweigerer unterscheiden sich in der Bewertung der Aussagen zu ihren grundlegenden Einstellungen und Verhaltensweisen – generelles Misstrauen, Extrovertiertheit und häufige Kommunikation mit anderen – nicht signifikant (siehe Tabelle 3). Das spricht dagegen, dass es sich bei diesen beiden Gruppen um unterschiedliche Menschentypen handelt. Dies gilt auch bei alleiniger Betrachtung der Erstkontakte.

Tabelle 1 Bewertung der letzten Umfrageerfahrung getrennt nach Teilnehmern und Verweigerern

Frage: „Denken Sie nun bitte an die letzte Befragung, bei der Sie mitgemacht haben. Was würden Sie sagen? War das eher eine angenehme oder eher eine unangenehme Erfahrung?“

	Eher angenehm (n=278)	Ambivalent (n=124)	Eher unangenehm (n=147)	Habe noch nie an einer Umfrage teilgenommen (n=54)	Alle Befragten (n=603)
	%	%	%	%	%
Teilnehmerinterviews (TI)	68	43	26	15	48
Verweigererinterviews (VI)	32	57	74	85	52
Summe	100	100	100	100	100

$Chi^2 = 98,6$ Cramer-V = ,404 $p < 0,001$

In Bezug auf die Einstellung zu und Erfahrung mit Umfragen zeigen sich (fast) durchgängig die erwarteten signifikanten Zusammenhänge mit der Teilnahmebereitschaft. Besonders die letzte Interviewerfahrung trennt die Teilnehmer von den Verweigerern: Die Befragten, die ihre letzte Befragung als angenehm empfanden, nehmen signifikant häufiger teil als diejenigen, die ihre letzte Befragung als unangenehm empfanden (siehe Tabelle 1). Auch bei getrennter Betrachtung der Panel- und der Erstkontakte zeigt sich dieses Ergebnis.

Die Daten sprechen auch für einen Oversurveying-Effekt, denn mit subjektiv empfundener Steigerung an Interviewanfragen in den vergangenen zwölf Monaten sinkt die Teilnahmebereitschaft deutlich und signifikant (siehe Tabelle 2). Zwar zeigen sich hier signifikante Unterschiede zwischen Panelkontakten und Erstkontakten – Erstere geben an, deutlich häufiger an Umfragen teilzunehmen –, allerdings besteht auch bei getrennter Betrachtung der Panel- und Erstkontakte in beiden Gruppen der erwartete Oversurveying-Effekt.

Tabelle 2 Unterschiede zwischen Teilnehmern und Verweigerern in Bezug auf Umfrageanfragen, Umfrageteilnahmen und Verhältnis von Teilnahmen zu Anfragen in den vergangenen 12 Monaten

Frage 1: „Wie oft sind Sie, abgesehen von der heutigen Befragung, in den letzten 12 Monaten gebeten worden, an einer Umfrage teilzunehmen?“

Frage 2 (falls bei Frage 1 eine Zahl genannt wurde): „Und wie oft haben Sie mitgemacht?“

	F	Eta	Signifikanz
Anzahl Umfrageanfragen in den vergangenen 12 Monaten	58,05	0,43	0,000
Anzahl Teilnahmen in den vergangenen 12 Monaten	13,67	0,44	0,000
Verhältnis Teilnahmen/Anfragen	151,72	0,56	0,000

Anova

Die generelle Bewertung von Umfragen und besonders persönliche Motive wie Spaß an Befragungen sind ebenfalls Unterscheidungskriterium zwischen Teilnehmern und Verweigerern. Altruistische Motive wie die Meinung, dass Umfragen wichtig und sinnvoll für die Gesellschaft sind, sind hingegen weniger entscheidend (siehe Tabelle 3). Das spricht dafür, dass potenziell Befragte sich nicht deshalb für die Teilnahme entscheiden, weil sie davon überzeugt sind, der Gesellschaft einen Dienst zu erweisen, sondern weil sie persönlich Interesse und Spaß daran haben. Es zeigen sich wenige Zusammenhänge mit soziodemografischen Merkmalen und grundlegenden Eigenschaften und Verhaltensweisen. Lediglich das generelle Misstrauen wirkt sich auf einige Aspekte der Umfrageeinstellung und die häufige Kommunikation mit anderen auf die Bewertung der letzten Interviewerfahrung aus. Da auch diese Zusammenhänge schwach sind, kann davon ausgegangen werden, dass das (indirekte) Verzerrungspotenzial dieser grundlegenden Persönlichkeitsmerkmale durch ihre Zusammenhänge mit der Umfrageeinstellung und -erfahrung gering ist. Dies gilt gleichermaßen für Panel- und Erstkontakte.

Merkmale der konkreten Befragung – Länge und Thema – zeigen keinen Zusammenhang mit der Teilnahmebereitschaft. Der Grund hierfür liegt möglicherweise in der fehlenden Eignung der Hauptbefragung für die Messung und der nicht möglichen Variation.¹¹ Allerdings kann die Tatsache, dass etwa ein Viertel aller Verweige-

11 Thema der Hauptbefragung war „Fragen zu Themen des täglichen Lebens“, die Länge wurde vom Interviewer in der Einleitung mit ca. 15 Minuten angekündigt. Beides liefert wenig Differenzierungspotenzial und konnte nicht variiert werden.

rer¹² am wesentlich kürzeren Verweigererinterview (mit dem entsprechend anderen Thema), das von einer anderen ‚Institution‘ durchgeführt wurde, teilnahmen, als Indiz dafür gewertet werden, dass die konkreten Befragungsmerkmale einen Einfluss haben, der anhand des Studiendesigns nur nicht gemessen werden konnte.

Tabelle 3 Unterschiede in den grundlegenden Einstellungen und den Umfrageeinstellungen der Teilnehmer und Verweigerer
Frage: „Ich lese Ihnen nun einige Aussagen vor. Bitte sagen Sie mir jeweils, inwieweit Sie diesen zustimmen [...]“

	F	Eta	Signifikanz
Es gibt nur sehr wenige Menschen, auf die man sich verlassen kann	0,03	0,13	0,865
Ich sage meine Meinung immer und überall offen	2,43	0,13	0,119
Ich spreche OFT mit anderen Menschen über Themen des täglichen Lebens	1,03	0,20	0,310
Umfragen bringen Abwechslung und sind interessant	157,36	0,48	0,000
Marktforschungsunternehmen behandeln die Daten vertraulich	62,52	0,32	0,000
Bei Umfragen wird häufig etwas gefragt, was niemanden etwas angeht	27,58	0,30	0,000
Oft weiß man gar nicht, ob mit einer Umfrage nicht eigentlich etwas verkauft werden soll	20,18	0,28	0,000
Umfragen sind für die Gesellschaft wichtig und sinnvoll	37,49	0,27	0,000
Bei den meisten Umfragen ist von vorneherein klar, was rauskommt	1,19	0,19	0,276

Anova

Während sich Panel- und Erstkontakte nicht in ihrer Bewertung der Wichtigkeit der Länge für ihre Entscheidung zur Teilnahme oder Verweigerung unterscheiden, geben die Erstkontakte, die verweigert haben, häufiger an, dass das uninteressante Thema entscheidungsbeeinflussend war. Die Panelkontakte kennen bereits aus dem letzten Jahr den Inhalt der Befragung und können das Thema deshalb besser einordnen.

12 Der wahre Anteil derer, die zu diesem Thema zum Interview bereit gewesen wären, liegt vermutlich weitaus höher, da 671 Verweigerer gar nicht mehr erneut kontaktiert wurden.

Dementsprechend spielen für ihre Entscheidung andere Faktoren eine Rolle und nicht das ihnen bekannte Thema, zu dem sie bereits ein Interview geführt haben.

Sowohl die fehlende Zeit im Moment der Interviewanfrage als auch die Bewertung des vergangenen Tages als stressig hängen signifikant mit der Teilnahmebereitschaft zusammen: Drei Viertel aller Teilnehmer geben an, vor allem deshalb teilgenommen zu haben, weil sie gerade Zeit hatten. Umgekehrt geben über 60% der Verweigerer an, dass die fehlende Zeit einer der Hauptgründe für ihre Verweigerung war. Einzuwenden ist, dass die Zeit als ‚Verweigerungsheuristik‘ (Schnell 1997: 161) angesehen werden kann: Die potenziell Befragten schieben die fehlende Zeit nur als Ausrede vor, um das Gespräch möglichst schnell zu beenden, obwohl andere Gründe maßgebend sind. Die Daten zeigen allerdings, dass diejenigen, die vor allem aus Zeitgründen das Interview abgelehnt haben, insgesamt die Einstellungen zu und Erfahrungen mit Umfragen deutlich positiver bewerten als die anderen Verweigerergruppen. Das spricht dafür, dass Zeitmangel als Verweigerungsgrund auch inhaltliche Bedeutung hat und nicht nur vorgeschoben wird. Dies gilt gleichermaßen für Panel- und Erstkontakte.

Die Interviewermerkmale Geschlecht, Alter und Erfahrung zeigen keine stabil messbaren Zusammenhänge mit der Teilnahmebereitschaft. Dennoch kann davon ausgegangen werden, dass der Interviewer bedeutsamer ist, als es die Daten vermuten lassen. Viele Einflüsse – vor allem Stimme und Sprache, die nach bisherigen Ergebnissen aus der Literatur den Haupteinfluss des Interviewers darstellen – konnten aufgrund des Studiendesigns nicht umfassend gemessen werden, denn die Interviewer wurden nach für die Hauptbefragung entscheidenden Gesichtspunkten und nach ihrer Verfügbarkeit eingesetzt. So handelte es sich z. B. in Hinblick auf Alter und Erfahrung um eine sehr homogene Gruppe.

Die bisherigen Betrachtungen lassen nur Rückschlüsse auf bivariate Zusammenhänge zwischen den möglichen Einflussgrößen und der Teilnahmebereitschaft zu. Entscheidend für das Verständnis des Prozesses, der zu Teilnahme oder Verweigerung führt, sind darüber hinaus die Zusammenhänge und Wechselwirkungen zwischen den einzelnen Einflussgrößen. Ziel ist es, die Einflussfaktoren zu identifizieren, die als stabile – unter Kontrolle (möglichst) aller anderen relevanten Faktoren gültige – Einflussgrößen auf die Teilnahmeentscheidung wirken. Hierfür eignet sich eine Regressionsanalyse. Da die abhängige Variable dichotom ist – Teilnahme oder Verweigerung –, wird eine logistische Regression gerechnet (vgl. Backhaus u. a. 2006: 425ff.). Dabei wird schrittweise anhand des Modells der Hierarchie der Einflussgrößen auf die Teilnahmebereitschaft vorgegangen (siehe Abbildung 2). Zunächst werden die soziodemografischen Merkmale als ‚Grundstock‘ einbezogen (Schritt 0). Wie bereits aufgrund der bivariaten Betrachtung zu erwarten war, ha-

ben die soziodemografischen Merkmale nahezu keinen Einfluss auf die Teilnahmeentscheidung (Nagelkerkes $R^2 = 0,085$).

Entsprechend dem Modell der Hierarchie der Einflussgrößen auf die Teilnahmebereitschaft ist zunächst für den potenziell Befragten die Bewertung der situativen Gegebenheiten entscheidend. Sie wirkt vermutlich als ‚Filter‘; nur wenn sie positiv bewertet wird, besteht die Chance, dass sich der potenziell Befragte mit dem Anliegen des Interviewers auseinandersetzt. Dementsprechend ist zu erwarten, dass die Einbeziehung der situativen Faktoren im Schritt 1 eine deutliche Steigerung der Güte des Gesamtmodells bewirkt. Dies ist auch der Fall: Nagelkerkes R^2 steigt deutlich an und erreicht einen „akzeptablen“ (Backhaus u. a. 2006: 456) Wert von 0,362. Die situativen Gegebenheiten tragen also einen entscheidenden Anteil zur Teilnahmeentscheidung bei.

Im nächsten Schritt wird der Interviewer als mögliche Einflussgröße mit einbezogen. Wie schon aufgrund der bivariaten Befunde zu vermuten war, trägt der Interviewer nicht viel zusätzliche Erklärungskraft bei (Nagelkerkes R^2 wächst um 0,017).

Im dritten Schritt werden die unabhängig von der konkreten Situation vorhandenen Einstellungen und Verhaltensweisen einbezogen. Hierunter fallen sowohl grundlegende Einstellungen und Verhaltensweisen als auch die Einstellung zu und die Erfahrung mit Umfragen. Die Einbeziehung der grundlegenden Eigenschaften (Schritt 3a) bringt wie erwartet keine Verbesserung der Güte des Gesamtmodells (Nagelkerkes R^2 steigt um 0,004). Keine der möglichen Einflussgrößen – Vertrauen in die Mitmenschen, Extrovertiertheit und häufige Gespräche mit anderen – erreicht das Signifikanzniveau. Zusammen mit den Befunden zu den soziodemografischen Merkmalen deutet dies darauf hin, dass grundlegende Persönlichkeitsmerkmale keinen Einfluss auf die Teilnahmeentscheidung haben. Durch die Einbeziehung der umfragespezifischen Einstellungen und Erfahrungen (Schritt 3b) hingegen verbessert sich die Erklärungskraft des Modells deutlich und kann als „sehr gut“ (Backhaus u. a. 2006: 456; Nagelkerkes $R^2 = 0,639$) bezeichnet werden. Neben den situativen Gegebenheiten tragen diese also entscheidend zur Teilnahmebereitschaft bei. Auch die einzelnen Faktoren haben fast alle einen signifikanten Einfluss auf die Teilnahmeentscheidung.

Im vierten und letzten Schritt werden die Merkmale der konkreten Befragung, also der Hauptbefragung, einbezogen. Die Güte des Gesamtmodells bleibt nahezu unverändert (Nagelkerkes R^2 steigt um 0,003) und auch die beiden Faktoren Länge und Thema der Hauptbefragung zeigen keinen signifikanten Einfluss. Wie ausgeführt, wird der Einfluss dieser beiden Faktoren vermutlich unterschätzt, einmal wegen der ungenauen Messung und der fehlenden experimentellen Variation, aber auch wegen der fehlenden Polarisierung des Themas.

Tabelle 4 Der Einfluss von den untersuchten Faktoren auf die Teilnahmebereitschaft

	Schritt 0 Sozio	Schritt 1 Zeit/ Situation	Schritt 2 Inter- viewer	Schritt 3a Grundl. Einst.	Schritt 3b Umfrage- einstellung	Schritt 4 Konkrete Befragung
	b	b	b	b	b	b
Geschlecht (Ref.-Kat: Weiblich)						
Männlich	,562**	,515*	,558**	,560**	,829**	,851**
Alter (Ref.-Kat: 60 J. und älter)						
18 bis 29 Jahre	-,008	,463	,520	,550	,140	,189
30 bis 44 Jahre	,380	,730	,778*	,776*	,797	,855
45 bis 59 Jahre	-,092	,203	,268	,297	,587	,599
Familienstand (Ref.-Kat.: Verw.)						
Ledig	,529	,533	,534	,528	,518	,590
Verheiratet/Zusammenlebend	,466	,652	,673	,672	,233	,263
Getrennt lebend/Geschieden	,748	,761	,815	,800	,453	,496
HHgröße (Ref.-Kat.: Mehrpers.HH)						
Einpersonenhaushalt	,896**	1,010**	1,023**	1,047**	,856	,879
Zweipersonenhaushalt	,266	,361	,359	,355	,490	,487
Bildung (Ref.-Kat.: Hoch)						
Niedrig	,217	,285	,300	,360	-,040	-,082
Mittel	-,041	,053	,018	,040	-,098	-,057
Beruf (Ref.-Kat.: Nicht arbeitend)						
Arbeitend	,291	,183	,152	,149	,292	,230
Wirtschaft. Lage (Ref.-Kat.: Schlecht)						
Gut	-,509	-,335	-,359	-,390	-,691	-,625
Befriedigend	-,272	-,297	-,316	-,345	-,953*	-,922*
Stressfaktor		-,010	-,016	-,016	-,128	-,135
Zeitfaktor		,723**	,703**	,702**	,829**	,802**
Geschl. Int. (Ref.-Kat.: Weiblich)						
Männlich			,563	,545	,660	,658
Alter Interviewer			-,046	-,044	-,031	-,032
Erfahrung (Ref.-Kat.: Erfahrung)						
Keine Erfahrung			-,433	-,433	-,407	-,374
Selbsteinschätzung			,371*	,371*	,469*	,459*
Faktor Interviewer			,161	,162	-,052	-,014
Vertrauen in Mitmenschen				-,026	,038	,030
Extrovertiertheit				-,146	-,179	-,191
Kommunikation/Kompetenz				,055	-,001	-,003
Erf. mit letzter Umfrage (Ref.-Kat.: Noch nie teilen)						
Eher angenehm					1,562**	1,561**
Ambivalent					1,392*	1,440*
Eher unangenehm					,980	1,028
Anzahl Umfrageanfragen					-,154**	-,152**
Verhältnis Teiln./Anfragen					,560**	,581**
Abwechslung und Interesse					,510**	,524**
Vertraulichkeit					,294**	,303**
Privatsphäre					-,294**	-,280**
Verkaufsanrufe					-,054	-,065
Wichtig und sinnvoll					,205	,215
Klar, was rauskommt					,316**	,306*
Faktor Länge						,128
Faktor Thema						-,132
Modellkonstante	-1,006	-3,817**	-4,710**	-4,305**	-6,963**	-7,070**
Pseudo R ² (Nagelkerkes)	,085	,362	,379	,383	,639	,642

n=627

Binär-Logistische Regression (0=Verweigerung; 1=Teilnahme)

Regressions-Koeffizienten (b) mit *: Signifikanzniveau < 0,05; mit **: Signifikanzniveau < 0,01

Auch in der multivariaten Betrachtung ist es entscheidend, die Erstkontakte getrennt zu betrachten, denn es ist anzunehmen, dass die Mechanismen beim ersten Kontakt anders ablaufen als bei Folgekontakten. Hierbei soll nur auf die zentralen Unterschiede eingegangen werden. Die Bedeutsamkeit der einzelnen Schritte des Modells der Hierarchie der Einflussgrößen auf die Teilnahmebereitschaft ist bei den Erstkontakten gleich: Den größten Anteil an der Erklärung haben die situativen Gegebenheiten (Schritt 1; Nagelkerkes R^2 steigt um 0,288) und die Einstellung zu und Erfahrung mit Umfragen (Schritt 3b; Nagelkerkes R^2 steigt um 0,259).

Allerdings zeigen zum Teil andere Faktoren einen signifikanten Einfluss: Bei den soziodemografischen Merkmalen hat das Alter einen Einfluss: Gehört ein Befragter den mittleren Altersgruppen – 30 bis 44 und 45 bis 59 Jahre – an, ist die Teilnahmewahrscheinlichkeit im Vergleich zu der Gruppe ‚60 Jahre und älter‘ größer. Diese Tendenz zeigt sich auch im Vergleich zu den 18- bis 29-Jährigen, erreicht hier aber nicht das Signifikanzniveau. Damit bestätigt sich der in der Literatur häufig replizierte Befund, dass alte Menschen eher verweigern. Die Einstellungen zu und Erfahrungen mit Umfragen zeigen unter den Erstkontakten nahezu die gleichen signifikanten Einflüsse wie unter allen Befragten.

Zusammenfassend gilt: Viele der bivariat gefundenen Zusammenhänge konnten auch im multivariaten Modell bestätigt werden, vor allem, dass die situativen Gegebenheiten und die Einstellungen zu und Erfahrung mit Umfragen den größten Beitrag zur Teilnahmeentscheidung leisten. Besonders die Tatsache, ob der potenziell Befragte im Moment der Interviewanfrage Zeit hat oder nicht und die Erfahrungen mit der letzten Befragung beeinflussen die Teilnahmebereitschaft. Das Gesamtmodell unter Einbeziehung aller Schritte weist zudem eine sehr gute Erklärungskraft auf. Das spricht dafür, dass die Mehrzahl der relevanten Faktoren identifiziert wurde.

6 Zusammenfassung und Bewertung der Ergebnisse

Die Befunde zeigen sowohl für bestimmte Traits als auch für bestimmte States einen Einfluss auf die Teilnahmebereitschaft, die jedoch im vorliegenden Fall allesamt nicht mit grundlegenden Persönlichkeitsmerkmalen zusammenhängen, was auf den ersten Blick die Folgerung nahelegt, dass in der Regel gravierende Verzerrungen für Umfragen aufgrund der Teilnehmerselbstselektion nicht zu befürchten sind. Als eine Haupteinflussgröße auf die Teilnahmebereitschaft haben sich hier die Erfahrungen mit und die Einstellung zu Umfragen erwiesen. Damit kann die Umfrageforschung durch gezielte Maßnahmen wie Imageverbesserung, kontinuierliche Öffentlichkeits-

arbeit, gut gewählte Einleitungstexte, vertrauensbildende Maßnahmen etc. Einfluss nehmen und so die Ausschöpfungsquoten erhöhen. Thema und Länge haben vermutlich einen – hier nicht gemessenen – Einfluss auf die Ergebnisse einer konkreten Befragung. Verzerrungen der Umfrageergebnisse im Allgemeinen sind jedoch auch dann nicht zu befürchten. Situative Faktoren sind einer der Haupteinflussfaktoren auf die Teilnahmebereitschaft. Sie zeigen schwache Zusammenhänge mit grundlegenden Persönlichkeitseigenschaften und bergen so ein geringes Verzerrungspotenzial. Insgesamt spricht dies im vorliegenden Fall für einen stark zufallsgesteuerten Ausfallprozess. Konvertierungsversuche zu einem anderen Zeitpunkt, eventuell auch durch einen anderen Interviewer, erscheinen somit Erfolg versprechend.

Insgesamt ergeben sich innerhalb unserer Studie somit kaum Hinweise darauf, dass Teilnehmer und Verweigerer grundlegend verschiedene Typen sind in Bezug auf soziodemografische Eigenschaften und ihre basalen Persönlichkeitseigenschaften. Im Fall der hier beispielhaft untersuchten Befragungsstudie sind somit inhaltliche Verzerrungen der Umfrageergebnisse durch Teilnahmeverweigerung unwahrscheinlich. Natürlich kann dieser Befund derzeit noch nicht auf andere Befragungen oder Umfragen allgemein generalisiert werden. Aus praktischen wie finanziellen Gründen konnte die Untersuchung nur im Rahmen einer einzelnen Marktforschungsstudie erfolgen; der Einfluss von Thema und Länge der konkreten Befragung oder des Interviewers konnte somit nicht durch echte Variation überprüft, sondern nur anhand der Selbsteinschätzung des Befragten als Größe einbezogen werden. Der Einfluss der Erhebungsart und des Auftraggebers konnte gar nicht untersucht werden. Es ist somit nicht auszuschließen, dass es hier zu Wechselwirkungen von Befragungs- und Erklärungsvariablen kommt, so dass es von Eigenschaften der Befragung wie Thema, Länge, Auftraggeber usw. abhängt, welche Befragtenmerkmale denn nun Verzerrungen durch Nonresponse hervorrufen. So könnte beispielsweise unser Befund, wonach Befragte mit positiver Umfrageeinstellung eine größere Bereitschaft zur Befragungsteilnahme aufweisen, bei bestimmten Befragungsthemen zu einem bedeutsamen Nonresponse Bias führen, nämlich immer dann, wenn das in der Umfrage zu erhebende Merkmal mit diesen Umfrageeinstellungen zusammenhängt, wie dies zum Beispiel im Bereich der politischen Partizipation für die Bereitschaft zur Teilnahme an politischen Wahlen dokumentiert ist. Aus dieser Perspektive kann somit nur von einer eingeschränkten Generalisierbarkeit der Befunde gesprochen werden. Deshalb ist weitere Forschung in diesem Bereich wichtig, die den Einfluss von Merkmalen der konkreten Befragung verlässlicher messen und vor allem über verschiedene Befragungsstudien hinweg vergleichen kann.

Zu guter Letzt gilt es natürlich noch auf die erkenntnistheoretische Paradoxie hinzuweisen, die unserem Untersuchungsdesign innewohnt: Auch diese Un-

tersuchung leidet – wie vergleichbare Nonresponse-Studien – unter Nonresponse. Nicht alle Verweigerer haben – aus verschiedenen Gründen – das Verweigererinterview geführt. Über die, die dies nicht getan haben, ist keine Aussage möglich. Es ist somit denkbar, dass auch unsere Ergebnisse durch einen Nonresponse-Bias verzerrt sind. Doch solange die Teilnahme an empirischer Sozialforschung auf Freiwilligkeit basiert und – gottlob – nicht erzwungen werden kann, wird diese verbleibende Unschärfe nie ganz auszuräumen sein.

Literatur

- Backhaus, K., B. Erichson, W. Plinke und R. Weiber, 2006: *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin/Heidelberg/New York: Springer.
- Biel, A. L., 1967: Abuses of Survey Research Techniques: The Phony Interview. *Public Opinion Quarterly* 31 (2): 298.
- Brosius, H.-B. und F. Koschel, 2001: *Methoden der empirischen Kommunikationsforschung. Eine Einführung*. Wiesbaden: Westdeutscher Verlag.
- De Leeuw, E. D. und W. De Heer, 2002: Trends in Household Survey Nonresponse. A Longitudinal and International Comparison. S. 41-54 in: R. M. Groves, D. A. Dillman, J. L. Eltinge und R. J. A. Little (Hg.): *Survey Nonresponse*. New York: Wiley.
- DeMaio, T. J., 1980: Refusals: Who, Where, and Why. *Public Opinion Quarterly* 44 (2): 223-233.
- Diekmann, A., 2006: *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Hamburg: Rowohlt Taschenbuch Verlag.
- Erbslöh, B. und A. Koch, 1988: Die Non-Response-Studie zum ALLBUS 1986. Problemstellung, Design, erste Ergebnisse. *ZUMA-Nachrichten* 22: 29-44.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_22.pdf (26.11.2008).
- Esser, H., 1973: Kooperation und Verweigerung beim Interview. S. 69-142 in: E. Erbslöh, H. Esser, W. Reschka und D. Schöne (Hg.): *Studien zum Interview*. Meisenheim am Glan: Verlag A. Hain.
- Esser, H., 1986: Über die Teilnahme an Befragungen. *ZUMA-Nachrichten* 18: 38-47.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_18.pdf (26.11.2008).
- Forsa, 2006: Akzeptanz von Interviews. Telefonmarketing: Überwiegend als Belästigung erlebt. *Context* 08/06: 2-6.
- Goyder, J., 1986: Surveys on Surveys: Limitations and Potentialities. *Public Opinion Quarterly* 50 (1): 27-41.
- Goyder, J., 1987: *The Silent Minority. Nonrespondents on Sample Surveys*. Cambridge: Polity Press.
- Groves, R. H., R. B. Cialdini und M. P. Couper, 1992: Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly* 56 (4): 475-495.
- Groves, R. M. und M. P. Couper, 1998: *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Groves, R. M. und L. E. Lyberg, 2001: An Overview of Nonresponse Issues in Telephone Surveys. S. 191-211 in: R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II und J. Waksberg (Hg.): *Telephone Survey Methodology*. New York: Wiley & Sons.
- Groves, R. H., S. Presser und S. Dipko, 2004: The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly* 68 (1): 2-31.

- Häder, S. und S. Gabler, 1998: Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. S. 9-18 in: S. Gabler, S. Häder und J. H. P. Hoffmeyer-Zlotnik (Hg.): Telefonstichproben in Deutschland. Opladen: Westdeutscher Verlag.
- Hansen, K. M., 2006: The Effect of Incentives, Interview Length, and Interviewer Characteristics on Response Rates in a CATI-Study. *International Journal of Public Opinion Research*. <http://ijpor.oxfordjournals.org/cgi/reprint/edl022v1> (15.05.07).
- Hartmann, E. L., H. L. Isaacson und C. M. Jurgell, 1968: Public Reaction to Public Opinion Surveying. *Public Opinion Quarterly* 32 (2): 295-298.
- Kaase, M. (Hg.), 1999: Qualitätskriterien der Umfrageforschung. Denkschrift Deutsche Forschungsgemeinschaft. Berlin: Akademie Verlag.
- Meier, G., M. Schneid, Y. Stegeman und A. Stiegler, 2005: Steigerung der Ausschöpfungsquote von Telefonumfragen durch geschickte Einleitungstexte. *ZUMA-Nachrichten* 57: 37-55. www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_57.pdf (9.12.2008)
- Oksenberg, L. und C. Cannell, 2001: Effects of Interviewer Vocal Characteristics on Nonresponse. S. 257-269 in: R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II und J. Waksberg (Hg.): *Telephone Survey Methodology*. New York: Wiley.
- Porst, R., 1998: Im Vorfeld der Befragung. Planung, Fragebogenentwicklung, Pretesting. *ZUMA-Arbeitsbericht* 98 (2). http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/zuma_arbeitsberichte/98_02.pdf (26.11.2008).
- Porst, R. und C. von Briel, 1995: Wären Sie vielleicht bereit, sich gegebenenfalls noch einmal befragen zu lassen? Oder: Gründe für die Teilnahme an Panelbefragungen. *ZUMA-Arbeitsbericht* 95 (4).
- Reuband, K.-H. und J. Blasius, 2000: Situative Bedingungen des Interviews, Kooperationsverhalten und Sozialprofil konvertierter Verweigerer. Ein Vergleich von telefonischen und face-to-face Befragungen. S. 139-170 in: V. Hüfken (Hg.): *Methoden in Telefonumfragen*. Wiesbaden: Westdeutscher Verlag.
- Rösch, G., 1998: Maßnahmen zur Reduktion von Stichprobenfehlern bei telefonischen Bevölkerungsumfragen. S. 9-18 in: S. Gabler, S. Häder, J. H. P. Hoffmeyer-Zlotnik (Hg.): *Telefonstichproben in Deutschland*. Opladen: Westdeutscher Verlag.
- Schleifer, S., 1986: Trends in Attitudes Toward and Participation in Survey Research. *Public Opinion Quarterly* 50 (1): 17-26.
- Schnell, R., 1997: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske+Budrich.
- Schräpler, J.-P., 2000: Was kann man am Beispiel des SOEP bezüglich Nonresponse lernen? *ZUMA-Nachrichten* 46: 117-150. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_46.pdf (26.11.2008).
- Sheets, T., A. Radlinski, J. Kohne und G. A. Brunner, 1974: Deceived Respondents: Once Bitten, Twice Shy. *Public Opinion Quarterly* 38 (2): 261-263.
- Singer, E., M. R. Frankel und M. B. Glassman, 1983: The Effect of Interviewer Characteristics and Expectations on Response. *Public Opinion Quarterly* 47 (1): 68-83.
- Smith, T. W., 1983: The Hidden 25 Percent. An Analysis of Nonresponse on the 1980 General Social Survey. *Public Opinion Quarterly* 47 (3): 386-404.
- Smith, T. W., 1984: Estimating Nonresponse Bias with Temporary Refusals. *Sociological Perspectives* 27 (4): 473-489.
- Smith, T. W., 2002: Developing Nonresponse Standards. S. 27-40 in: R. M. Groves, D. A. Dillman, J. L. Eltinge, J. A. L. Roderick (Hg.): *Survey Nonresponse*. New York: Wiley.
- Stocké, V. und B. Becker, 2004: Determinanten und Konsequenzen der Umfrageeinstellung. Bewertungsdimensionen unterschiedlicher Umfragesponsoren und die Antwortbereitschaft der Befragten. *ZUMA-Nachrichten* 54: 89-116. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_54.pdf (26.11.2008).

Stocké, V. und B. Langfeldt, 2003: Umfrageklima in Deutschland. *Context* 14/03: 6–8.

Tuckel, P. und H. O'Neill, 2002: The vanishing Respondent in Telephone Surveys. *Journal of Advertising Research* September–October 2002: 26–48.

Van der Vaart, W., Y. Ongena, A. Hoogendroom und W. Dijkstra, 2005: Do Interviewers' Voice Characteristics influence Cooperation Rates in Telephone Surveys? *International Journal of Public Opinion Research* 18 (4):488–499.

Korrespondenzadressen: Gregor Daschmann
Institut für Publizistik
Colonel-Kleinmann-Weg 2
D-55099 Mainz
gregor.daschmann@uni-mainz.de

Anna Schnauber
Kellerstraße 22
D-65183 Wiesbaden
anna.schnauber@gmx.de

Der Einfluss von Kontrollüberzeugungen der Interviewer auf die Teilnahme an telefonischen Befragungen

The Influence of Interviewer's Locus of Control on the Participation in Telephone Interviews

*Christian Weidmann, Patrick Schmich
und Sabine Schiller-Born*

Zusammenfassung

Um Unterschiede in der Überzeugungskraft von Interviewern in telefonischen Befragungen besser zu verstehen, wird der Erklärungsbeitrag von Kontrollüberzeugungen der Interviewer für die Teilnahmebereitschaft von potentiellen Interviewpartnern an Befragungen bestimmt. Studien zu Kontrollüberzeugungen aus anderen Berufszweigen lassen die Annahme zu, dass Interviewer mit einer starken internalen Kontrollüberzeugung, die also das eigene Handeln und die gewählten Strategien als ausschlaggebend für Erfolg und Misserfolg erachten, höhere Erfolgsquoten aufweisen. Diese These wird anhand von Daten des am Robert Koch-Institut erhobenen Telefonischen Gesundheitssurveys 2007 überprüft. Wie sich in logistischen Mehrebenenregressionen zeigt, haben Interviewer mit hoher internaler Kontrollüberzeugung tatsächlich höhere Erfolgsquoten. Kontrollüberzeugungen der Interviewer und die damit verbundenen Nutzenerwartungen hinsichtlich der eigenen Bemühungen können somit einen eigenständigen Beitrag leisten, um die Unterschiede in den Erfolgsquoten der Interviewer zu erklären. Wie sich weiter zeigt, sind Kontrollüberzeugungen allerdings kein geeignetes Kriterium, um Interviewer für Anrufe mit einer schwierigen Ausgangsbasis auszuwählen.

Abstract

In order to reach a better understanding of the differences in persuasive power of interviewers conducting telephone polls this explanatory contribution is dealing with the locus of control of these interviewers and the rate of their success at persuading potential interviewees to take part in such surveys. Studies on the locus of control in other professions assume an increase in the rate of success for interviewers with a strong internal locus of control.

This hypothesis is tested on the basis of data from the Telephone Health Survey 2007 conducted by the Robert Koch-Institute. Results from a logistic multilevel model show that interviewers with high internal locus of control really have higher success rates. These results underline the influence that the convictions and the expectations of the interviewers have on the rate of participation in telephone interviews. However, locus of control does not appear to be a valid interviewer selection criterion when the goal is to convince target persons who are averse to surveys.

1 Einleitung¹

Telefonisch erhobene Befragungen spielen in der Sozial- und Gesundheitsforschung auf Grund der vergleichsweise geringen Kosten und der erleichterten Feldorganisation eine zunehmend wichtige Rolle, sind jedoch mehr als andere Befragungen mit der Problematik sinkender Responsequoten konfrontiert (vgl. Curtin/Presser/Singer 2005). Die steigende Anzahl der telefonisch durchgeführten Befragungen – der Anteil der Telefoninterviews in der Markt- und Meinungsforschung hat sich seit 1990 mehr als verdoppelt² – und die Zunahme von Werbeanrufen haben die Bereitschaft zur Teilnahme an telefonischen Befragungen in der Bevölkerung erheblich sinken lassen. Mit der steigenden Skepsis wächst die Bedeutung der Person des Interviewers³ für die Realisierung eines Interviews, da in den ersten Sekunden eines telefonischen Kontaktes die Seriosität und die Bedeutung der durchgeführten Studie vermittelt werden muss. Die erfolgreiche Überwindung dieser Kontaktpphase bis hin zur ersten inhaltlichen Frage, in der die Mehrzahl der Verweigerungen ausgesprochen wird, ist in der Regel die schwierigste Herausforderung im Rahmen der Interviewertätigkeit (vgl. Hüfken/Schäfer 2003). Überzeugungskraft, Durchsetzungsvermögen und ein seriöses Auftreten zählen daher zu den Qualitätsmerkmalen, die einen guten Interviewer charakterisieren und die eine hohe Ausschöpfungsquote sichern.

Um die zum Teil erhebliche Varianz in den Erfolgsquoten der Interviewer zu erklären, wurden bislang zahlreiche Einflussgrößen auf der Interviewerebene diskutiert. Erfahrung in der Interviewertätigkeit wurde hierbei wiederholt als positiver Faktor genannt, da erfahrene Interviewer in der Regel höhere Erfolgsquoten aufweisen (vgl. Hansen 2006; Couper/Groves 1992; Schmich 2006) und die Interviewdauer niedriger ausfiel (vgl. Olson/Peytchev 2007). Die Befunde von Hüfken/Schäfer (2003) legen allerdings nahe, dass entscheidend sein könnte, wo diese Erfahrung gesammelt wurde. Einsätze in Marktforschungsunternehmen etwa, in denen die Bedeutung der Erfolgsquote hinter der Anzahl der realisierten Interviews zurück tritt, könnten sich negativ auf den Überzeugungserfolg auswirken.

- 1 Wir danken I. Deideck für die Unterstützung bei der Erstellung des Manuskripts sowie S. Kienzle, F. Tschan, A. Loerbroks, S. Schneider und den beiden anonymen Gutachtern für zahlreiche wertvolle Hinweise.
- 2 Laut ADM hat sich der Anteil von 22% im Jahr 1990 auf 46% im Jahr 2006 erhöht (vgl. ADM, 2006: Jahresbericht 2006. Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. http://www.adm-ev.de/pdf/Jahresbericht_06.pdf (01.09.2008)).
- 3 Wenn im Folgenden von Interviewern gesprochen wird, sind stets Interviewerinnen und Interviewer gemeint.

Positiv wirkte sich indessen ein fortgeschrittenes Lebensalter der Interviewer aus (vgl. Hansen 2006; Hox/de Leeuw 2002; Schmich 2006). In einigen Studien konnten weibliche Interviewer höhere Überzeugungserfolge erzielen, wobei hier zum Teil widersprüchliche Befunde vorliegen (vgl. Hansen 2006). Stimmliche Merkmale wurden ebenso diskutiert wie das persönliche Auftreten in der Interviewsituation (vgl. Oksenberg/Coleman/Cannell 1986; Hüfken/Schäfer 2003). In der Studie von Hüfken/Schäfer (2003) kam einem enthusiastischen Auftreten des Interviewers die mit Abstand größte Bedeutung zu, es überragte alle weiteren Interviewermerkmale hinsichtlich des Erklärungsbeitrags.

In weiteren Studien wurden auch Einstellungen der Interviewer zu ihrer Tätigkeit erfasst und untersucht. Hierbei erwiesen sich Einstellungen zum Umgang mit ablehnenden Zielpersonen als ausschlaggebend. Waren die Interviewer der Meinung, Verweigerungen seien zu akzeptieren und die Privatsphäre der Zielpersonen zu schützen, fielen die Responsequoten niedriger aus als bei Interviewern, die der Meinung waren, jeder sei von einer Teilnahme zu überzeugen und Überzeugungsbemühungen bei ablehnenden Probanden seien prinzipiell akzeptabel (de Leeuw/Hox/Snijkers/de Heer 1998; Hox/de Leeuw 2002). Weniger eindeutig belegt ist der Einfluss der generellen Erwartungshaltung der Interviewer hinsichtlich ihres Einflusses auf das erfolgreiche Zustandekommen eines Interviews. Einerseits berichten Couper und Groves (1992) tatsächlich von höheren Erfolgsraten bei jenen Interviewern, die der Meinung sind, sie können jede auch noch so ablehnende Zielperson von einer Teilnahme überzeugen. Andererseits erwies sich in der Studie von Link (2006) eine allzu optimistische Einstellung zur Tätigkeit des Interviewers und den Möglichkeiten der Probanden-Gewinnung zu Beginn einer Interviewer-Karriere als eher negativ für die spätere Effizienz als Interviewer.

In beiden Studien wurde die Einschätzung des individuellen Einflusses auf den Interviewerfolg jedoch nur am Rande und wenig differenziert erfasst, wohingegen eine umfassende Analyse der generellen Kontrollüberzeugungen von Interviewern im Hinblick auf den Erfolg ihrer Überzeugungsbemühungen bislang noch aussteht. Aus anderen Berufsfeldern ist jedoch schon seit längerer Zeit bekannt, dass eine interne Kontrollüberzeugung – also die Vorstellung, dass der Schlüssel zu Erfolg und Misserfolg vor allem in den eigenen Händen liegt – zu besserer beruflicher Performanz führt als externe Kontrollüberzeugungen (vgl. Judge/Bono 2001). Internale Kontrollüberzeugungen von Interviewern in Bezug auf ihre Tätigkeit als Interviewer sollten daher ebenfalls mit größerem Überzeugungserfolg verbunden sein und einen eigenständigen Erklärungsbeitrag zum Verständnis variierender Erfolgsquoten der Interviewer liefern. Unseres Wissens nach existiert bislang jedoch keine Studie, in der Kontrollüberzeugungen der Interviewer angemessen berücksichtigt wurden.

Ziel dieses Artikels ist es daher, die These eines Zusammenhangs zwischen Kontrollüberzeugungen und Erfolg bzw. Misserfolg bei der Gewinnung von Teilnehmern in telefonischen Befragungen empirisch zu überprüfen, um Unterschiede in den Erfolgsquoten der Interviewer besser verstehen zu können. Hierzu wird zunächst im zweiten Abschnitt auf den theoretischen Zusammenhang zwischen Kontrollüberzeugungen und Interviewerfolg eingegangen, woraus drei zu überprüfende Hypothesen entwickelt werden. Im dritten Abschnitt werden der Telefonische Gesundheitssurvey 2007, der die Datengrundlage unserer empirischen Analyse bildet sowie die verwendeten Operationalisierungen und das Auswertungskonzept vorgestellt. Im vierten Abschnitt folgt die Darstellung der empirischen Ergebnisse im Hinblick auf die zuvor formulierten Hypothesen. Der fünfte Abschnitt schließt mit einer Diskussion der Befunde sowie der daraus folgenden Konsequenzen für die Analyse von Interviewereffekten und den Einsatz von Interviewern in telefonischen Befragungen.

2 Kontrollüberzeugung und Interviewerfolg

Aufbauend auf der Beobachtung, dass sich Wahrnehmungsmuster der Zielerreichung stark unterscheiden und durch die zugeschriebene Bedeutung der eigenen Bemühungen klar abgrenzen lassen, wurde von amerikanischen Psychologen in den 50er und 60er Jahren des 20. Jahrhunderts das Konzept des ‚Locus of Control‘ entwickelt (vgl. Rotter 1966). Dieser Ansatz kennt zwei idealtypische Kontrollüberzeugungsmuster: während bei einer internalen Kontrollüberzeugung das eigene Handeln und die zugrunde liegenden Strategien für das Erreichen eines bestimmten Ziels verantwortlich gemacht werden, sind bei externalen Kontrollüberzeugungen außen stehende Kräfte maßgeblich. Nicht das eigene Handeln, sondern einflussreiche Dritte, die Umstände oder schlicht Glück führen zum Erreichen oder Verfehlen eines Ziels. Der Ansatz des ‚Locus of Control‘ fokussiert dabei nicht die spezifische Reaktion in konkreten Situationen, sondern die generalisierte Erwartungshaltung jedes Einzelnen im Hinblick auf das Lösen von Problemlagen.

Über die generelle Unterscheidung zwischen internalen und externalen Kontroll-Überzeugungen hinaus wurde das Konzept des ‚Locus of Control‘ auf zahlreiche Verhaltensbereiche angewandt (vgl. Furnham/Steele 1993). Insbesondere Gesundheitsverhaltensmuster konnten mit internalen und externalen Gesundheits-Kontrollüberzeugungen in Verbindung gebracht werden, wobei internale Überzeugungen eher mit einem gesundheitsförderlichen Verhalten einhergingen (vgl. Steptoe/Wardle 2001). Im Bereich des Berufslebens konnten mit Skalen zum ‚Work

Locus of Control' ebenfalls Zusammenhänge hergestellt werden. So erwiesen sich Arbeitnehmer mit einer hohen internalen Kontrollüberzeugung am Arbeitsplatz nicht nur als zufriedener als ihre weniger internal ausgerichteten Kollegen, sondern hatten auch eine bessere berufliche Performanz (vgl. Judge/Bono 2001). Als Ursache für die höhere Arbeitsplatzzufriedenheit wurden stärkeres Engagement, höhere Entlohnungen in Folge größerer Erfolge, frühere Beförderungen und die geringere Neigung, in unbefriedigenden Beschäftigungsverhältnissen zu verharren, genannt (vgl. Spector 1982; Muhonen/Torkelson 2004). Die bessere berufliche Performanz wurde vor allem mit den verinnerlichten Wahrnehmungsmustern und Nutzenerwartungen begründet. Im Falle dominanter externaler Kontrollüberzeugungen erscheint das eigene Bemühen nur mit geringem Nutzen verbunden, da externe Kräfte über Erfolg und Misserfolg entscheiden. Besondere Gratifikationen für das eigene Streben werden nicht oder nur kaum erwartet. Anders im Falle stark ausgeprägter internaler Überzeugungen. Hier liegt die Erwartungshaltung bezüglich des Nutzens eigener Anstrengungen deutlich höher, was im Falle ansprechender Entlohnung zu einem deutlich stärkeren Engagement führt (vgl. Judge/Bono 2001).

Auch bei der Tätigkeit als Interviewer ist vor diesem Hintergrund eine positive Wirkung internaler Kontrollüberzeugungen zu erwarten. Interviewer sind angehalten, möglichst viele Zielpersonen von der Teilnahme am Interview zu überzeugen und Verweigerungen möglichst zu verhindern, da die Ausschöpfung in telefonischen Befragungen ein zentrales Qualitätskriterium darstellt. Das Konzept des ‚Locus of Control‘ legt hierbei nahe, dass die Überzeugungsbemühungen der Interviewer um so engagierter ausfallen werden, je stärker sie damit rechnen, dass ihr eigenes Handeln und ihr Auftreten in der Kontaktphase maßgeblichen Einfluss auf das Zustandekommen des Interviews hat. Sind sie dagegen eher der Überzeugung, dass es von den Inhalten der Studie, von der Laune des Befragten oder schlicht vom Zufall und somit nicht von ihren eigenen Strategien abhängt, werden sich Interviewer weniger bemühen und in der Konsequenz weniger erfolgreiche Kontaktierungsbemühungen aufweisen können. Voraussetzung ist allerdings, dass auch die Interviewer ein Interesse haben, eine hohe Erfolgsquote zu erzielen. Dies trifft in der Regel zu, da Interviewer von der Feldleitung entsprechende Vorgaben erhalten, deren Einhaltung überprüft wird. Des Weiteren werden Interviewer oftmals zumindest in Teilen leistungsbezogen bezahlt, so dass eine Anreizstruktur gegeben sein sollte.⁴ So kann eine erste Hypothese zum Einfluss von Kontrollüberzeugungen auf die Interviewer-Tätigkeit formuliert werden.

4 Auch die Entlohnung im Telefonischen Gesundheitssurvey 2007 erfolgte in Abhängigkeit von der realisierten Interviewzahl. Vgl. hierzu Abschnitt 3.1.

Hypothese 1: Interviewer mit hoher internaler Kontrollüberzeugung verfügen über größere Nutzen-Erwartungen hinsichtlich ihrer Überzeugungs-Bemühungen und weisen daher ein stärkeres Engagement und in der Folge einen höheren Anteil an realisierten Interviews auf, als Interviewer mit niedriger internaler Kontrollüberzeugung. Hohe externale Kontrollüberzeugungen wirken sich entsprechend negativ auf die Erfolgsquote aus.

Anrufe, die in einer Interviewteilnahme oder einer Verweigerung enden, haben sehr oft eine Vorgeschichte, da der erste Kontakt nicht immer zu einer Entscheidung führt. In vielen Fällen kann zunächst nur mit einem Haushaltsmitglied (sog. Kontaktperson) und nicht mit der gesuchten Person (Zielperson) gesprochen werden oder aber, die Zielperson ist zwar anwesend, aber aus zeitlichen oder sonstigen Gründen nicht bereit teilzunehmen. Der Verlauf dieser vorherigen Anrufe lässt in vielen Fällen gute Prognosen für die Erfolgsaussichten eines erneuten Anrufs zu (vgl. Hüfken/Schäfer 2003). Wurde von einer Kontaktperson bereits eine stellvertretende Verweigerung ausgesprochen, oder die Zielperson nannte lediglich einen vagen Termin für einen Anruf („...rufen Sie mich doch nächste Woche noch mal an“), stehen die Chancen eines Interviews in der Regel wesentlich schlechter, als wenn mit der Zielperson ein genauer Termin vereinbart werden konnte und ein prinzipielles Teilnahmeinteresse geäußert wurde. Interviewer können oftmals die Historie dieser früheren Anrufe einsehen, ehe ein erneuter Anruf erfolgt. Auch bei den Surveys des Robert Koch-Instituts stehen den Interviewern die Notizen zu den früheren Anrufen zur Verfügung, was oftmals eine Erwartungshaltung auf Seiten der Interviewer entstehen lässt.

Liefert die Kontakthistorie Anzeichen für eine ablehnende Haltung der Zielperson, werden die Kontrollüberzeugungen der Interviewer vermutlich zu besonders deutlichen Unterschieden in den Erfolgsquoten führen. Interviewer mit externalen Kontrollüberzeugungen werden voraussichtlich schneller aufgeben, da sie ihren Gestaltungsspielraum als geringer einschätzen. Interviewer mit internalen Kontrollüberzeugungen werden sich dagegen von einer vergleichsweise schlechten Ausgangsbasis weniger entmutigen lassen, da sie primär ihre eigene Strategie und ihre Bemühungen als maßgeblich erachten und die Ausgangsbasis somit in den Hintergrund tritt. Die Folge dieser unterschiedlichen Reaktionen auf negative Vorzeichen dürfte ein besonders starker Einfluss der Kontrollüberzeugung in einer schlechten Ausgangssituation sein. Wir erwarten daher in unseren Analysen eine Interaktion von Kontrollüberzeugung und Ausgangsbasis. Die Bestätigung dieser Erwartung hätte für die Steuerung des Interviewereinsatzes erhebliche praktische Implikationen, da Kontrollüberzeugungen der Interviewer dann als Steuerungskriterium für die Zuordnung von schwierigen Anrufen dienen könnten. In Hypothese 2 ist die These einer Interaktion zusammengefasst.

Hypothese 2: Können die Interviewer aus der Kontakthistorie schließen, dass die zu befragende Person einem Interview eher ablehnend gegenüber steht, ist der Einfluss der Kontrollüberzeugungen besonders groß. Interviewer mit internaler Kontrollüberzeugung lassen sich von einer schlechten Ausgangsbasis weitaus weniger entmutigen und haben daher deutlich höhere Erfolgsquoten.

Bislang wurde von einer einseitigen kausalen Beeinflussung des Interviewerfolgs durch die vorliegende Kontrollüberzeugung gesprochen. Diese Interpretationsweise entspricht dem eingangs vorgestellten Verständnis von Kontrollüberzeugungen als grundlegende und von einzelnen Problemen losgelöste Wahrnehmungsmuster. Im Bereich des ‚Work Locus of Control‘ erscheinen jedoch auch Rückwirkungen der Arbeitserfahrungen auf die Wahrnehmungsmuster als möglich. So konnten beispielsweise Andrisani und Nestel (1976) mit Panel-Analysen von Bevölkerungsbefragungen den positiven Einfluss von Beförderungen, Gehaltserhöhungen und dem Wiedereintritt in den Arbeitsmarkt nach vorheriger Arbeitslosigkeit auf die Entwicklung einer internalen Kontrollüberzeugung belegen. Offensichtlich haben hier die gesammelten Erfahrungen im Erwerbsalltag ihre Spuren in den Wahrnehmungsmustern hinterlassen. In einer der wenigen konzeptionellen Arbeiten zu diesem Thema sprechen Furnham und Steele (1993) von Beeinflussungsspiralen, in denen sich, ausgehend von wiederholt positiven oder negativen Erfahrungen, Zuversicht und Motivation langfristig verändern können und auch die grundlegenden Kontrollüberzeugungen eine Wandlung erfahren. In einem selbstverstärkenden Prozess beeinflussen diese gewandelten Kontrollüberzeugungen schließen wiederum Motivation, Engagement und letztlich die neuen Erfahrungen von Erfolg und Misserfolg. Muhonen und Turkolsen (2004) erwarten daran anknüpfend einen positiven Einfluss von Führungspositionen auf die Entwicklung einer internalen Kontrollüberzeugung von Führungskräften. Untersuchungen in dieser spezifischen Gruppe scheinen die Vermutung zu unterstützen (vgl. Siu/Spector/Cooper/Donald 2001). Über den Ablauf und den zeitlichen Rahmen von Veränderungen der Kontrollüberzeugungen und die Rückwirkung auf den beruflichen Erfolg ist bislang aber nur wenig bekannt.

Für die Entwicklung von Kontrollüberzeugungen und Interviewerfolg in Telefonbefragungen erscheint uns die von Furnham und Stelle (1993) beschriebene Beeinflussungsspirale als ein plausibles Modell. Da Interviewer bei Verweigerungen durchaus verbal angegriffen werden können und allgemein den Erfolg ihrer Bemühungen unmittelbar erfahren, ist vor allem in längeren Interviewerkarrieren eine Anpassung der grundlegenden Wahrnehmungsmuster naheliegend. In diesem Sinne sollten positive Rückmeldungen eine optimistische Sichtweise und interne Kontrollüberzeugungen begünstigen, die in der Konsequenz zu einem stärkeren

Engagement und damit verbunden zu größeren Kontaktierungserfolgen führen. Gleichzeitig können negative Rückmeldungen und ein als gering empfundener Handlungsspielraum pessimistischere Einstellungen und externale Kontrollüberzeugungen begünstigen. Auf Grund des geringeren Engagements sind wiederum eher negative Rückmeldungen zu warten, die eine erneut negative Beeinflussung der Kontrollüberzeugungen verursachen können.

Als Konsequenz dieser beiden entgegengesetzten Spiralen sollte die Bedeutung der Kontrollüberzeugung für den Interviewerfolg in einem erfahrenen Interviewerstamm deutlicher hervortreten, da sich internale Kontrollüberzeugungen und Erfolg sowie externale Überzeugungen und Misserfolg im Laufe einer Interviewerkarriere gegenseitig verstärken. Entsprechend erwarten wir in einer abschließenden Hypothese eine Interaktion von Interviewererfahrung und Kontrollüberzeugung. Unter erfahrenen Interviewern sollte auf Grund der beschriebenen Beeinflussungsspiralen im Laufe ihrer bisherigen Interviewertätigkeit Kontrollüberzeugung und Interviewerfolg enger miteinander verknüpft sein.

Hypothese 3: Unter erfahrenen Interviewern tritt der Einfluss von Kontrollüberzeugungen auf den Interviewerfolg eher stärker hervor, da sich Interviewerfolg und Kontrollüberzeugungen im Laufe einer Interviewerkarriere gegenseitig verstärken.

3 Methodik

Die empirische Überprüfung des Zusammenhangs zwischen Kontrollüberzeugungen der Interviewer und den erzielten Erfolgen bei der Gewinnung von Teilnehmern erfolgte mit Hilfe des Telefonischen Gesundheitssurveys 2007, der vom Robert Koch-Institut in Berlin erhoben wurde. Im Folgenden soll diese Befragung kurz vorgestellt werden, ehe die Operationalisierung zentraler Merkmale und das Auswertungskonzept beschrieben wird.

3.1 Der Telefonische Gesundheitssurvey 2007

Im Rahmen der Entwicklung eines Gesundheitsmonitorings werden am Robert Koch-Institut schon seit 2003 regelmäßig repräsentative telefonische Gesundheitssurveys erhoben, die Auskunft geben sollen über die Verbreitung chronischer Krankheiten und zentraler Risikofaktoren in der deutschen Bevölkerung sowie über die Inanspruchnahme von Leistungen des Gesundheitswesens (vgl. Kohler/Rieck/Borch 2005). Auf Grund der geringen Feldkosten und der raschen Verfügbarkeit

der Daten stellen die telefonischen Gesundheitssurveys eine flexible Ergänzung zu den in größeren Abständen erhobenen Untersuchungssurveys des Robert Koch-Instituts dar.

Der Telefonische Gesundheitssurvey 2007 wurde vom 19. Januar bis zum 17. März 2007 im CATI (Computer unterstütztes Telefoninterview) -Labor des Robert Koch-Instituts durchgeführt. Er diente vor allem der Weiterentwicklung und Verbesserung der eingesetzten Instrumente. Mit einem eigenen Interviewerstamm konnten 2.503 standardisierte Interviews realisiert werden. Es wurden Fragen zum Gesundheitsverhalten, zu chronischen Erkrankungen, zu körperlichen Beeinträchtigungen im alltäglichen Leben, zur Inanspruchnahme medizinischer Leistungen und zur psychischen Gesundheit gestellt.

Die Stichprobe wurde aus der erwachsenen, deutschsprachigen Wohnbevölkerung in Privathaushalten mit Festnetzanschlüssen auf Basis des Gabler-Häder-Designs gezogen (vgl. Gabler/Häder 2002). GESIS stellte hierzu einen Auswahlrahmen zu Verfügung, der auf dem Gesamtbestand aller öffentlich zugänglichen Telefonnummern aufbaut. Um sicher zu stellen, dass auch der wachsende Anteil von nicht eingetragenen Nummern berücksichtigt wird, wurden von GESIS nach der Bereinigung von offensichtlichen Sondernummern und Geschäftsanschlüssen die beiden letzten Ziffern aller bekannten Telefonnummern abgeschnitten und durch die Ziffern von 0 bis 99 ergänzt. Aus diesem Auswahlrahmen wurden 22.497 Nummern uneingeschränkt zufällig ausgewählt und angerufen. Sofern es sich um eine gültige Nummer innerhalb eines Privathaushaltes handelte, wurde versucht, mit dem erwachsenen Haushaltsmitglied, das zuletzt Geburtstag hatte, ein Interview zu realisieren. Die zweite Auswahlstufe in Form der ‚Last-Birthday-Methode‘ sollte sicher stellen, dass alle Mitglieder von Mehrpersonenhaushalten die gleiche Auswahlchance haben, unabhängig von der Erreichbarkeit im Haushalt.

Wie Tabelle 1 zeigt, konnte bei 7.969 Rufnummern über einen persönlichen Kontakt mit einer Kontaktperson⁵ sicher gestellt werden, dass die Rufnummer vergeben ist. In 5.868 Fällen konnte schließlich eine Zielperson über die Geburtstagsfrage eindeutig ermittelt und auch kontaktiert werden. Aus diesen Zielpersonenkontakten konnten während der Feldzeit 2.503 vollständige Interviews durchgeführt werden. Zusätzlich konnten 499 Verweigerer zu Kurzinterviews bewegt werden, in denen einige zentrale inhaltliche Fragen sowie Angaben zu Gründen der Nicht-Teilnahme erhoben wurden.

5 Kontaktpersonen können im Haushalt lebende Personen sein, wohingegen die Zielperson jene Person ist, die als letztes Geburtstag hatte und die interviewt werden soll. Anders als etwa bei Befragungen auf der Basis von Einwohnermeldestichproben ist die Ermittlung der Zielperson im Rahmen des Gabler-Häder-Designs erst während der Feldarbeit möglich.

Tabelle 1 Response-Übersicht des Telefonischen Gesundheitssurveys 2007 (vgl. Kohler/Rieck/Borch 2005)

	Absolut N	Brutto- stichprobe	Netto- stichprobe	Personen- kontakte	Zielpersonen- kontakte
Bruttoansatz	22.497	100,0			
1 Telekom-Ansage / Falsche Nummer	10.096	44,9			
2 Max-Kontakt erreicht	302	1,3			
3 FAX / Modem	630	2,8			
4 Kein Privathaushalt	793	3,5			
5 ZP im Haushalt unbekannt		0,0			
6 Verstorben, verzogen	13	0,1			
7 Sprachprobleme	162	0,7			
8 Sonstige neutrale Ausfälle	21	0,1			
Neutrale Ausfälle insgesamt (1-8)	12.017	53,4			
Bereinigtes Brutto	10.480	46,6	100,0		
9 Teilnehmer nimmt nicht ab	2.019	9,0	19,3		
10 Anschluss besetzt	17	0,1	0,2		
11 Anrufbeantworter	475	2,1	4,5		
Personenkontakte insgesamt	7.969	35,4	76,0	100,0	
12 Terminvereinbarung nicht wahrnehmbar	98	0,4	0,9	1,2	
13 Kein Termin möglich während Feldzeit	130	0,6	1,2	1,6	
14 KP ist beschäftigt	0	0,0	0,0	0,0	
15 KP gibt keine Auskunft: Thema, Projekt	0	0,0	0,0	0,0	
16 KP gibt keine Auskunft: anderer Grund	139	0,6	1,3	1,7	
17 KP verweigert absolut jede Teilnahme	1.734	7,7	16,5	21,8	
Zielpersonenkontakte insgesamt	5.868	26,1	56,0	73,6	100,0
18 ZP ist beschäftigt	939	4,2	9,0	11,8	16,0
19 ZP gibt keine Auskunft: Thema, Projekt	0	0,0	0,0	0,0	0,0
20 ZP gibt keine Auskunft: anderer Grund	12	0,1	0,1	0,2	0,2
21 ZP verweigert absolut jede Teilnahme	1.824	8,1	17,4	22,9	31,1
22 ZP ist krank, nicht in der Lage	35	0,2	0,3	0,4	0,6
23 Abbruch	29	0,1	0,3	0,4	0,5
24 Interview nicht auswertbar	27	0,1	0,3	0,3	0,5
Systematische Ausfälle insgesamt (9-24)	7.478	33,2	71,4		
Durchgeführte Interviews	3.002	13,3	28,6	37,7	51,2
- davon Vollinterviews	2.503	11,1	23,9	31,4	42,7
- davon Kurzfragebögen	499	2,2	4,8	6,3	8,5

KP = Kontaktperson

ZP = Zielperson

Die Ausschöpfungsquote, also der Anteil realisierter Interviews an allen Zielpersonenkontakten, lag bei 42,7%. Berücksichtigt man zusätzlich die Kurzfragebögen, so lag die generelle Kooperationsrate⁶ der identifizierten Zielpersonen bei 51,2%. Diese Quoten liegen niedriger als in den früheren telefonischen Gesundheitssurveys des Robert Koch-Instituts (vgl. Kohler/Rieck/Borch 2005), wobei hinzu gefügt werden muss, dass der inhaltliche Schwerpunkt des Telefonischen Gesundheitssurvey 2007 in erster Linie auf dem Testen der eingesetzten Instrumente lag und die Feldzeit vergleichsweise kurz ausfiel. Angesichts des hohen Anteils offener Terminvereinbarungen ist zu vermuten, dass bei einer längeren Feldzeit die Responsequote noch deutlich zu steigern gewesen wäre.

Während der Feldzeit kamen 41 Interviewer zum Einsatz, wovon 24 Interviewer schon über Erfahrungen aus mindestens einem früheren telefonischen Gesundheitssurvey am Robert Koch-Institut verfügten. Die Interviewer unterteilten sich in 23 Frauen und 18 Männer und die Altersspanne reichte von 21 bis 49 Jahren, wobei die Mehrheit (24 Interviewer) das 30. Lebensjahr noch nicht erreicht hatte. In einer dreitägigen Schulung im Vorfeld der Erhebung wurde den Interviewern unter anderem der Inhalt der Studie, der Umgang mit der eingesetzten Erhebungssoftware, das Vorgehen in der Kontaktierungsphase sowie das grundsätzliche Auftreten im Interview vermittelt. Die Bezahlung der Interviewer setzte sich aus einem fixen Grundbetrag pro Arbeitstag sowie einer Pauschale für jedes realisierte Interview zusammen.

Telefoniert wurde von Montag bis Freitag zwischen 16.00 und 20.00 Uhr sowie an Samstagen zwischen 14.00 und 18.00 Uhr, wobei pro Arbeitstag maximal 20 Interviewer arbeiten konnten. Die Felddauer umfasste 49 Arbeitstage, an denen im Schnitt pro eingesetztem Interviewer 3,4 Interviews realisiert werden konnten. Die durchschnittliche Interviewzeit lag bei 32 Minuten.

Die individuellen Erfolgsquoten der Interviewer, also das Verhältnis von erfolgreichen Interviews (ohne Kurzinterviews) zu dokumentierten Verweigerungen, weist eine große Spannweite auf. Während im Schnitt das Verhältnis bei 0,7 Verweigerungen pro durchgeführtem Interview lag, gab es einerseits fünf Interviewer, die mehr Verweigerungen als Interviews protokollierten und andererseits einen Interviewer, der gar keine Verweigerung ausgesprochen bekam sowie weitere sechs Interviewer mit weniger als 0,5 Verweigerungen pro Interview. Die Interviewer wurden von drei Supervisoren betreut, die für Rückfragen zur Verfügung standen und die bei jedem Interviewer mindestens zwei Interviews mitverfolgten, um

6 Im Anschluss an die frühere Dokumentation der Telefonischen Gesundheitssurveys des Robert Koch-Instituts verwenden wir den Begriff Kooperationsquote als Anteil der Voll- und Kurzinterviews an allen Zielpersonenkontakten (vgl. Kohler/Rieck/Borch 2005).

die Einhaltung der Qualitätsstandards zu sichern. Die Durchführung der Interviews erfolgte als Computer unterstütztes Telefoninterview (CATI) mit Hilfe der Software ‚Voxco Interviewersuite‘. Die Zuweisung der Nummern erfolgte zufällig durch die eingesetzte Software und entsprechend der bereits in früheren telefonischen Gesundheitssurveys entwickelten Anrufsteuerung (vgl. Kohler 2005). Wurde eine Nummer kontaktiert ohne dass ein Interview durchgeführt werden konnte, waren die Interviewer angehalten einen kurzen Kommentar zum Verlauf der Kontaktierung einzugeben. Bei einer späteren Kontaktierung standen den Interviewern diese Informationen dann zur Verfügung.

3.2 Operationalisierungen

Die Anzahl der realisierten Interviews und die dabei verbuchten Verweigerungen durch die Zielpersonen geben unmittelbar Auskunft über die Überzeugungskraft der Interviewer, sofern unterstellt werden kann, dass sich schwer bzw. leicht zu überzeugende Zielpersonen gleich über die Interviewer verteilen. Angesichts der zufälligen Nummerzuspielung durch die Interviewsoftware ist von einer solchen Gleichverteilung tatsächlich auszugehen. Für die Analysen wurden daher alle jene Fälle herangezogen, in denen entweder ein komplettes Interview durchgeführt wurde⁷ (n=2.503, siehe Tabelle 1) oder eine grundsätzliche Verweigerung vorlag (n=1.824). Diese Verweigerungen wurden von den Zielpersonen selbst und nicht stellvertretend durch ein Haushaltsmitglied geäußert. Stellvertretende Verweigerungen durch Kontaktpersonen wurden bei den Berechnungen nicht berücksichtigt, da die zugehörigen Nummern im Laufe des Feldes ein zweites Mal kontaktiert wurden, um nach Möglichkeit mit der Zielperson selbst zu sprechen.

Nicht ausgeschlossen werden kann jedoch, dass einige faktische Verweigerungen durch Zielpersonen von den Interviewern nicht als solche kodiert, sondern mit anderen Ergebniscodes oder einer Terminvereinbarung versehen wurden, um die eigene Quote nicht zu verschlechtern. Stichprobenartige Kontrollen durch die eingesetzten Supervisoren ergaben hierfür allerdings keine Anhaltspunkte.

Für die Erfassung der Kontrollüberzeugungen der Interviewer standen neben Skalen zur generellen Kontrollüberzeugung einige Skalen zum ‚Work Locus of Control‘ bereit (vgl. Spector 1988; Dubois 1997 Furnham/Steele 1993), auf deren Einsatz aber nach Rücksprache mit den Interviewern verzichtet wurde. Da die Interviewertätigkeit eine spezialisierte und im Arbeitsablauf stark regulierte Aufgabe

7 Die durchgeführten Kurzinterviews wurden in den Analysen nicht berücksichtigt.

mit geringen Freiheitsgraden ist, die in der Regel nicht hauptberuflich ausgeübt wird, stießen zahlreiche Items zum beruflichen Alltag bei den Interviewern auf Unverständnis. Insbesondere Fragen zum individuellen Einfluss auf die Entlohnung und zu Aufstiegschancen erschienen vor dem Hintergrund der Interviewertätigkeit als unpassend. Die Interviewer wurden daher gebeten, vier Items zu beantworten, die von uns direkt auf die Interviewertätigkeit und die Überzeugung von Zielpersonen zugeschnitten wurden und die Kontrollüberzeugungen hinsichtlich des Interviewerfolgs erfassen sollten. Erhoben wurden diese Items während des Feldverlaufs im Vorfeld einer Schicht, wobei jeder Interviewer bereits einige Einsätze hatte, ehe der Fragebogen ausgefüllt wurde. Der jeweilige Feldtag des Ausfüllens lässt sich jedoch leider nicht mehr rekonstruieren.

Wie in Tabelle 2 abgebildet, wurden je zwei Items formuliert, die eine interne bzw. eine externe Kontrollüberzeugung ausdrücken. Zur Auswahl standen vier Antwortkategorien⁸, wobei auf eine mittlere Kategorie, die weder Zustimmung noch Ablehnung ausdrückt, bewusst verzichtet wurde. In anderen Kontexten wurde bei Fragen, die auf externe Einstellungen zielen, von hohen Zustimmungswerten in diesen mittleren Kategorien berichtet, was möglicherweise auf sozial erwünschtes Antwortverhalten schließen lässt und die Zusammenfassung der Items erschwert (vgl. Jakoby/Jacob 1999). In einer Faktorenanalyse wurde die Zuordnung unserer Items zu zwei unabhängigen Faktoren schließlich bestätigt (siehe Tabelle 2).⁹

Für die weiteren Analysen wurden die vier Items analog zum Vorgehen bei Rotter (1966) zu einem Index addiert, wobei ein hoher Skalenwert (Maximalwert=16) in der vorliegenden Analyse eine stark interne Kontrollüberzeugung und ein niedriger Wert (Minimalwert=4) eine stark externe Kontrollüberzeugung zum Ausdruck bringen. Die Index-Werte der Interviewer streuten zwischen 6 und 13 Punkten, hierbei lag der Median bei 11 und das arithmetische Mittel bei 9,8 Punkten. Ein einzelner Interviewer verweigerte die Teilnahme an der Befragung, so dass lediglich die Interviews der 40 teilnehmenden Interviewer in die Analysen eingehen konnten.

Um indirekte Einflüsse auszuschließen und den reinen Erklärungsbeitrag von Kontrollüberzeugungen einschätzen zu können, wurden für die Analysen weitere bekannte Einflussgrößen des Interviewerfolgs als Kontrollvariablen herangezogen.

8 Diese lauteten: 1 ‚lehne voll und ganz ab‘, 2 ‚lehne eher ab‘, 3 ‚stimme eher zu‘ und 4 ‚stimme voll und ganz zu‘. Für die spätere Index-Bildung wurden die Antwortvorgaben der externen Items in ihrer Reihenfolge umcodiert.

9 Detaillierte Auswertungen des Antwortverhaltens bestätigten die Trennschärfe der Items. Lediglich einer der vierzig Interviewer hat den internalen Aussagen ebenso deutlich zugestimmt, wie den externalen.

Tabelle 2 Items zur Messung der tätigkeitsbezogenen Kontrollüberzeugungen

Items	Kontroll- über- zeugung	Zustimmung ¹ (in Prozent)				Faktor 1 (Eigenwert 1,697) ²	Faktor 2 (Eigenwert 1,109) ²
		lehne voll ab	lehne eher ab	stimme eher zu	stimme voll zu		
Es liegt vor allem an mir selbst, ob sich Befragte zur Teilnahme motivieren lassen.	Intern	0	38,5	53,8	7,7	0,808	-0,326
Prinzipiell kann jeder zur Teilnahme an einer Befragung bewegt werden.	Intern	35,0	32,5	25,0	7,5	0,903	0,096
Das Zustandekommen eines Interviews hängt hauptsächlich vom Inhalt der Studie ab.	Extern	2,5	35,0	55,0	7,5	0,036	0,825
Ob jemand an einer Befragung teilnimmt, ist eine Frage zufälliger Ereignisse.	Extern	5,0	57,5	35,0	2,5	-0,181	0,713

1 Die Prozentzahlen beziehen sich auf die 40 Interviewer, die an Befragung teilgenommen haben. Beim ersten Item machte ein Interviewer keine Angaben. Die Prozentangaben beziehen sich hier auf die 39 gültigen Angaben.

2 Ergebnisse einer Faktorenanalyse (Principal Component Analyse mit Varimax-Rotation) zur Überprüfung der Itemzugehörigkeit. Bei den Werten handelt es sich um die jeweiligen Faktorladungen.

Als zusätzliche Interviewercharakteristika gingen Geschlecht und das Alter in Jahren in die Analysen ein. Des Weiteren wurde die Anzahl früherer Studien am Robert Koch-Institut, an dem der jeweilige Interviewer mitgearbeitet hat, als Indikator für die Erfahrung als Interviewer verwendet¹⁰. Unberücksichtigt blieb dagegen die Erfahrung als Interviewer in anderen Instituten, da entsprechende Informationen nicht vorlagen. Neben Interviewercharakteristika wurden zwei Feldmerkmale als Kontrollvariablen berücksichtigt. Um auszuschließen, dass Veränderungen im Feldverlauf oder Unterschiede in Häufigkeit und Verlauf früherer Kontaktierungen den Einfluss von Kontrollüberzeugungen verdecken oder überlagern, wurden der jeweilige Feldtag und die Kontaktierungsart in die späteren Analysen einbezogen. Die Feldtage wurden hierzu durchnummeriert und gingen fortlaufend in die Berechnungen ein.

Bei den Kontaktierungen wurde unterschieden, wie die Ausgangsbasis für den Anruf, der zu einer Teilnahme oder Verweigerung führte, einzuschätzen ist. Konnte der Interviewer oder einer seiner Kollegen in einem früheren Anruf bereits mit der Zielperson sprechen, wurde eine gute Ausgangsbasis unterstellt. Von einer strikten

10 Die Teilnahme an Studien-Pretests wurde hierbei als separate Studie gewertet, da diese in der Regel mit einer eigenen Schulung und einer intensiven Feldarbeit verbunden waren.

Ablehnungshaltung kann bei einem Zweitanruf nicht mehr ausgegangen werden, da ein solcher erneuter Anruf nur erfolgt, wenn die Zielperson zuvor nicht verweigerte.

Wurde dagegen in einem früheren Anruf von einer Kontaktperson eine stellvertretende Verweigerung ausgesprochen, ist von einer schlechten Ausgangsbasis auszugehen. Zwar kann nicht generell von der Haltung einer Kontaktperson auf die der Zielperson geschlossen werden, eine gewisse Homogenität lässt sich jedoch erwarten. Da bei stellvertretenden Verweigerungen auf jeden Fall ein erneuter Anruf erfolgte, kann für diese Anrufe daher von einer schlechten Ausgangsbasis gesprochen werden. Zur Erfassung der beiden Ausgangssituationen wurde jeweils eine Dummy-Variable gebildet.

3.3 Auswertungskonzept

Die Überprüfung des Einflusses von Kontrollüberzeugungen der Interviewer auf die Teilnahme von Zielpersonen an der Befragung erfolgte mit Hilfe eines logistischen Regressionsmodells. Da die 1.824 Verweigerungen und 2.503 Teilnahmen von lediglich 41 Interviewern erhoben wurden und die Daten somit eine hierarchische Struktur aufweisen, wurde hierbei auf die Methode der Mehrebenenanalyse mit Hilfe der Software MLwiN 2.02 zurückgegriffen (vgl. Rasbach/Steele/Browne/Prosser 2005). Die untere Ebene bildete in den Analysen der einzelne Anruf, der durch den Feldtag und die Ausgangssituation weiter charakterisiert wurde. Die zweite Ebene verkörperte der Interviewer und dessen Charakteristika. In die Berechnung gingen auf dieser Ebene das Alter und das Geschlecht des Interviewers, seine Erfahrung und die Kontrollüberzeugungen ein. Mit Hilfe des Wald-Tests wurden alle geschätzten Koeffizienten auf Signifikanz überprüft.

Die abhängige Variable hatte im Modell die Ausprägung 1 ‚Zielperson hat teilgenommen‘ und 0 ‚Zielperson hat verweigert‘. Zur Überprüfung von Hypothese 2 und 3 wurden abschließend Interaktionsterme von Kontrollüberzeugung und Kontaktierungsart sowie von Kontrollüberzeugung und Interviewererfahrung eingebracht.

Der Achsenabschnitt wurde hierbei als variabel über die Interviewer hinweg modelliert, wodurch den unterschiedlichen Fähigkeiten der Interviewer bei der Gewinnung von Zielpersonen Rechnung getragen wird (vgl. Blohm/Hox/Koch 2006). Des Weiteren wurde jeweils getestet, ob die erklärenden Merkmale in ihrer Bedeutung signifikant zwischen den Interviewern variieren und somit als Zufallskoeffizient zu berücksichtigen sind. Da für logistische Mehrebenenmodelle keine Log-Likelihood-Statistik zur Verfügung steht, wurde hierzu ebenfalls der Wald-Test herangezogen (vgl. Twisk 2006). Bei keiner der erklärenden Merkmale zeigte sich

jedoch eine deutliche Varianz, so dass alle im Folgenden vorgestellten Modelle über einen variablen Achsenabschnitt und feste Koeffizienten verfügen.

4 Ergebnisse der Analyse

In Tabelle 3 sind die Odds Ratios sowie die 95-prozentigen Konfidenzintervalle der logistischen Mehrebenen-Regressionsmodelle abgebildet. Das erste Modell enthält zunächst keine erklärenden Variablen, sondern lediglich einen Achsenabschnitt. Dieser wurde als variabel über die Interviewer hinweg modelliert und reflektiert die unterschiedlichen Erfolgsquoten der Interviewer bei der Gewinnung von Zielpersonen. Die Varianz dieses Achsenabschnitts liegt bei 22% mit einem Standardfehler von 6%.

Das zweite Modell enthält zusätzlich zum Achsenabschnitt die Merkmale der Anrufebene. Wie deutlich wird, hat der Feldtag des Anrufs keinen nennenswerten Einfluss auf das Zustandekommen des Interviews. Das Odds Ratio liegt hier genau bei 1,00. Ein Zuwachs der Verweigerungen im Laufe des Feldes, wie von Hüfken/Schäfer (2003) berichtet, ist somit nicht zu verzeichnen. Zwar ist auch beim Telefonischen Gesundheitssurvey 2007 die absolute Zahl der realisierten Interviews gegen Feldende gesunken. Dies war allerdings in der vermehrten Anwahl von solchen Nummern begründet, unter denen niemand erreicht werden konnte und die noch bis zur vorher definierten Mindestanzahl kontaktiert werden mussten. Wie die vorgestellten Analysen aus Tabelle 3 belegen, hat sich hierbei das Verhältnis von erfolgreichen Anrufen zu Verweigerungen jedoch nicht verändert.

Einen sehr starken Einfluss hat dagegen die Ausgangssituation. Kann diese als schlecht beschrieben werden, da in einem vorherigen Anruf eine stellvertretende Verweigerung geäußert wurde, sinkt die Chance auf eine Interviewrealisierung im erneuten Anruf laut Modell 2 auf ein Viertel gegenüber einer neutralen Ausgangsbasis (Odds Ratio von 0,25). Dieser höchst signifikante Effekt lässt sich auch bereits losgelöst vom Regressionsmodell in der reinen Verteilung der Anrufe erkennen. So konnten 406 Anrufe mit einer solch schlechten Ausgangsbasis registriert werden, wovon lediglich 100 (24,6%) zu einem realisierten Interview führten.

Tabelle 3 Logistisches Multilevel-Regressionsmodell zur Erklärung der Teilnahme am Interview

Variablen	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5
	Null-Modell	Anrufmerkmale	Interviewermerkmale	Gesamtmodell	Interaktionsmodell
	Odds Ratio (95-prozentiges Konfidenz- intervall)				
Fixed					
<i>Anrufebene</i>					
Feldtag		1,00 (1,00-1,00)	1,00 (1,00-1,00)	1,00 (1,00-1,00)	1,00 (1,00-1,00)
Kontaktierungsart					
- Neutral		Referenz	Referenz	Referenz	Referenz
- schlecht		0,25 (0,19-0,31)	0,25 (0,19-0,31)	0,25 (0,19-0,31)	0,25 (0,19-0,31)
- gut		2,29 (1,92-2,74)	2,30 (1,94-2,73)	2,31 (1,94-2,76)	2,29 (1,92-2,74)
<i>Interviewerebene</i>					
Alter			1,02 (1,00-1,05)	1,03 (1,01-1,05)	1,03 (1,01-1,05)
Geschlecht					
- Männer			Referenz	Referenz	Referenz
- Frauen			1,22 (0,90-1,64)	1,17 (0,87-1,57)	1,14 (0,87-1,50)
Erfahrung			1,05 (0,99-1,12)	1,05 (0,99-1,11)	0,82 (0,58-1,17)
Kontroll-Index				1,08 (1,02-1,15)	1,03 (0,93-1,14)
Erfahrung*Kontroll-Index					1,03 (0,99-1,07)
Random					
Intercept	1,50 (1,28-1,75)	1,46 (1,20-1,78)	0,61 (0,28-1,34)	0,27 (0,09-0,81)	0,40 (0,12-1,32)
N ^{Anrufebene}	4296	4296	4296	4296	4296
N ^{Interviewerebene}	40	40	40	40	40
Varianz (Std.fehler) ^{Intercept}	0,22 (0,06)	0,21 (0,06)	0,16 (0,05)	0,14 (0,04)	0,12 (0,04)
Veränderung der Varianz	Referenz	-4,5%	-27,3%	-36,4%	-45,4%

Anders ist das Verhältnis, wenn bereits zuvor mit der Zielperson gesprochen werden konnte und somit eine gute Ausgangsbasis besteht. 689 der insgesamt 915 Interviews mit einer guten Ausgangsbasis endeten mit einem Interview, was einer Quote von 75% entspricht.¹¹ Das zugehörige Odds Ratio aus Modell 2 verdeutlicht, dass dementsprechend die Chancen für ein Interview in einer solch guten Ausgangsbasis 2,29-mal höher sind als bei einem neutralen Anruf ohne vorherigen Zielpersonkontakt. Die Ausgangsbasis ist demnach ein wichtiger Einflussfaktor, der erklären kann, warum einzelne Anrufe eher zum Erfolg führen als andere.

Wenig Aufschluss bietet die Ausgangsbasis dagegen für das Verständnis der variierenden Erfolgsquoten der Interviewer. Die Varianz des Achsenabschnitts sinkt in Modell 2 nur marginal auf 0,21. Es ist also nicht davon auszugehen, dass die unterschiedlichen Erfolgsquoten darauf zurück zu führen sind, dass die Anrufe mit guter und schlechter Ausgangsbasis ungleich über die Interviewer verteilt sind. In einem separaten Modell mit variierendem Koeffizienten konnte zudem – wie unter 3.3 berichtet – keine nennenswerte Varianz der beiden Koeffizienten über die Interviewer ausgemacht werden. Anhaltspunkte, dass die Interviewer mit einer guten oder schlechten Ausgangsbasis unterschiedlich umgehen konnten, liegen demnach nicht vor. Entsprechend konnte auch in einem separaten, in Tabelle 3 nicht abgebildeten Modell, kein signifikanter Interaktionseffekt von Kontrollüberzeugung und Ausgangsbasis bestimmt werden. Interviewer mit internaler Kontrollüberzeugung konnten mit einer schlechten Ausgangsbasis nicht besser umgehen als ihre Kollegen mit externalen Kontrollüberzeugungen. Die Erwartungen aus Hypothese 2 wurden somit nicht bestätigt und Empfehlungen für einen gezielten Interviewereinsatz lassen sich, anders als erhofft, auf Basis unserer Analysen nicht ableiten.

In Modell 3 und 4 wurden im nächsten Schritt zusätzlich die Merkmale der Interviewer berücksichtigt, wobei die Kontrollüberzeugungen erst in Modell 4 aufgenommen wurden, um deren singulären Beitrag zur Reduzierung der Interviewer-Varianz bestimmen zu können. Wie sich zunächst zeigt, bleiben die Koeffizienten der Anrufebene durch die zusätzlichen Merkmale weitgehend unberührt. Die Effekte des Feldtags und der Ausgangssituation sind somit unabhängig von den Interviewermerkmalen. Auf der Ebene der Interviewermerkmale erweist sich neben dem Alter insbesondere die Kontrollüberzeugung der Interviewer als sehr einflussreich. Mit jedem weiteren Punkt auf der Kontrollskala in Richtung internaler Kontrollüberzeugung steigt die Chance ein Interview zu realisieren um acht Prozentpunkte (siehe Odds Ratio von 1,08 in Modell 4). Die erste der eingangs formulierten Hypothesen,

11 Aus den 3006 Anrufen mit einer neutralen Ausgangsbasis gingen 1714 (57,0%) Interviews hervor.

wonach sich eine interne Kontrollüberzeugung positiv auf die Teilnahme am Interview auswirkt, kann demnach bestätigt werden. Kontrollüberzeugungen können als signifikanter Einflussfaktor des Interviewerfolgs verstanden werden, der einen Teil der Varianz in den Erfolgsquoten erklärt. Wie aus Modell 4 ersichtlich wird, sinkt der Varianzanteil des variablen Achsenabschnitts durch Berücksichtigung der Kontrollüberzeugung um weitere 9 Prozentpunkte gegenüber Modell 3, das die übrigen Interviewermerkmale und die Charakteristika des Anrufs beinhaltet (-36,4 vs. -27,3). Kontrollüberzeugungen liefern demnach einen eigenständigen Erklärungsbeitrag zum Verständnis der unterschiedlichen Erfolgsquoten der Interviewer.

Auch das Alter des Interviewers hat einen positiven Einfluss auf das Zustandekommen eines Interviews. Wie in Modell 4 erkennbar, steigen die Chancen eines Interviews mit jedem Lebensjahr des Interviewers um 3% an. Dieser Befund stimmt mit den Ergebnissen von Hansen (2006), Hox/de Leeuw (2002) und Schmich (2006) überein, wobei in unseren Analysen erst unter Berücksichtigung der Kontrollüberzeugungen von einem statistisch signifikanten Alterseinfluss gesprochen werden kann (vgl. Modell 4 vs. Modell 3). Hierbei muss allerdings berücksichtigt werden, dass unsere Interviewer, wie berichtet, in der Mehrzahl das 30. Lebensjahr noch nicht vollendet hatten. Für detaillierte Analysen des Alterseffektes, der nicht im Mittelpunkt unserer Arbeit stand, sollte hier auf einen ausgeglichenen Interviewerstamm geachtet werden.

Kein signifikanter Effekt zeigte sich dagegen beim Erfahrungsniveau der Interviewer. Zwar ließen die Befunde in den Studien von Hansen (2006) sowie Couper/Groves (1992) deutliche Einflüsse der Erfahrung in der Interviewertätigkeit erwarten. Das Konfidenzintervall des Erfahrungswertes in Modell 3 und 4 widerspricht jedoch diesen Erwartungen. Interessanterweise fällt das Konfidenzintervall in einem logistischen Regressionsmodell ohne Berücksichtigung der hierarchischen Datenstruktur erheblich kleiner aus und lässt den Schluss eines positiven Erfahrungseinflusses zu. Möglicherweise wurde in bisherigen Studien auf Grund der mangelnden Berücksichtigung der hierarchischen Struktur der Erfahrungseffekt überschätzt. Ebenfalls nicht ausgeschlossen werden kann jedoch auch, dass die Zahl der 41 eingesetzten Interviewer zu gering war, um den Erfahrungswert deutlicher hervortreten zu lassen.

Auch hinsichtlich der Bedeutung der Kontrollüberzeugungen scheint die Erfahrung der Interviewer unerheblich. Wie aus Modell 5 ersichtlich wird, erweist sich ein Interaktionskoeffizient von Erfahrung und Kontrollüberzeugung als nicht signifikant. Die in Hypothese 3 geäußerte Erwartung, wonach der Kontrollüberzeugungseffekt unter erfahrenen Interviewern deutlicher erkennbar sei, muss daher zurück gewiesen werden. Die These einer Beeinflussungsspirale, im Rahmen derer sich interne Kontrollüberzeugungen und Interviewerfolg sowie externe Kon-

trollüberzeugungen und Misserfolg gegenseitig verstärken, kann auf der Basis der vorgestellten Analysen nicht bestätigt werden. Da der Interaktionskoeffizient den kritischen Signifikanzwert von $p < 0,05$ jedoch nur knapp verfehlt und die Varianz auf der Interviewerebene im Interaktionsmodell nochmals reduziert wird (-45,4% vs. -36,4% im Modell 4), besteht aber auch hier die Möglichkeit, dass die Zahl der eingesetzten Interviewer nicht ausreicht, um die These überprüfen zu können. Weitere Analysen stehen hier noch aus, insbesondere zu den Verzerrungen, die aus einem vorzeitigen Ausscheiden aus der Tätigkeit als Interviewer, resultieren können.

5 Diskussion

Die vorgestellten Ergebnisse belegen erstmals, dass Kontrollüberzeugungen der Interviewer einen wichtigen Beitrag zum Verständnis der zum Teil erheblichen Unterschiede in den Erfolgsquoten der Interviewer liefern können. Internale Kontrollüberzeugungen gehen mit deutlich höheren Erfolgsquoten in telefonischen Befragungen einher, wohingegen Interviewer mit externalen Kontrollüberzeugungen vergleichsweise niedrigere Quoten aufweisen.

Diese Relevanz der Kontrollüberzeugungen liefert zahlreiche Implikationen für die Praxis und für zukünftige Studien zum Einfluss von Interviewermerkmalen. In der Praxis können Kontrollüberzeugungen in der Rekrutierung neuer Interviewer eingesetzt und als Indikator für eine erfolgreiche Mitarbeiterentwicklung verwendet werden. Da Interviewer mit internalen Kontrollüberzeugungen besonders hohe Erfolgsquoten aufweisen, sollten die Institute vor dem Hintergrund sinkender Ausschöpfungsquoten Bedingungen schaffen, die internale Kontrollüberzeugungen und positivere Nutzen-Erwartungen der Bemühungen begünstigen. Entsprechende Schulungen zur Optimierung von Kontaktierungsstrategien, ein positives Arbeitsklima und Hilfestellungen für die Interviewer bei der Verarbeitung von übertrieben negativen Reaktionen mancher Zielpersonen erscheinen uns hierbei ebenso geeignet wie Beschäftigungsbedingungen, die den Interviewern eine längerfristige Perspektive bieten. Frustration und Fluktuation im Interviewerstamm könnten reduziert, die Ausschöpfungsquoten stabilisiert und auch die Gesundheit der Mitarbeiter geschützt werden. Da internale Kontrollüberzeugungen den Umgang mit beruflichem Stress erleichtern, ist eine Förderung internaler Kontrollüberzeugungen nicht nur lohnenswert im Hinblick auf die Erfolgsquoten, sondern auch ein Beitrag, die gesundheitlichen Risiken der fordernden Interviewer-Tätigkeit zu reduzieren (vgl. Muhonen/Torkelson 2004).

Anders als von uns erwartet sind Kontrollüberzeugungen dagegen nicht als Auswahlkriterium für die Bearbeitung von Anrufen mit einer schwierigen Ausgangsbasis geeignet. In unseren Analysen konnten keine relevanten Erfolgsunterschiede zwischen Interviewern mit internalen und externalen Kontrollüberzeugungen in schwierigen Kontexten ausgemacht werden. Auch die weiteren Interviewermerkmale wie Alter, Geschlecht und Erfahrung interagierten in separaten Analysen nicht mit der Ausgangsbasis. In der Praxis erscheinen die Erfolgsquoten der Interviewer daher als einzig verlässliches Kriterium um Interviewer für besonders schwierige Zielgruppen auszuwählen. Weitere explorative Analysen zum Interviewerfolg in einer schwierigen Ausgangsbasis sind also angezeigt.

Abgesehen von den praktischen Implikationen empfehlen die vorgestellten Befunde Kontrollüberzeugungen der Interviewer in zukünftigen Studien zum Einfluss von Interviewermerkmalen zu berücksichtigen. Die bisherigen Versuche Interviewererwartungen durch einzelne Items oder durch einen generellen Optimismus zu erfassen (vgl. Singer/Frankel/Glassman 1983; Couper/Groves 1992; Link 2006) sollten durch eine systematische Erhebung von Kontrollüberzeugungen ersetzt werden, da diese inhaltlich konsistent interpretiert werden können und die Gefahr von Messfehlern gegenüber einzelnen Items reduziert wird. Dass in unseren Analysen das Alter der Interviewer erst unter Berücksichtigung der Kontrollüberzeugungen als signifikanter Einflussfaktor hervorgetreten ist (vgl. Modell 4 vs. Modell 3), unterstreicht die Bedeutung von Kontrollüberzeugungen als möglichen Konfounder. Dies sollte in multivariaten Modellen zum Interviewerfolg berücksichtigt werden.

Darüber hinaus sollten künftige Studien zu Interviewereffekten unbedingt die hierarchische Datenstruktur beachten, die durch einen bei Telefonsurveys üblicherweise kleinen Interviewerstamm mit einer großen Anzahl realisierter Interviews entsteht. Wie berichtet, fielen in einem separaten Regressionsmodell ohne Berücksichtigung der Datenstruktur die Konfidenzintervalle der Interviewererfahrung erheblich kleiner aus. Diese wäre somit ohne Mehrebenenmodellierung als signifikanter Einflussfaktor ausgewiesen worden. Die große Bedeutung, die man der Interviewererfahrung bisher beigemessen hat, ist daher eher kritisch zu hinterfragen (vgl. Couper/Groves 1992; Hansen 2006).

Des Weiteren sollten Studien, die Kontrollüberzeugungen der Interviewer erfassen, einige Punkte beachten, die sich im Rahmen unserer Studie als einschränkend erwiesen haben. So sollte eine komplexere Erfassung von Kontrollüberzeugungen im Interviewerkontext angestrebt werden, die sich auf mehr als vier Items stützt und gleichzeitig einen grundlegenden Bezug zur spezifischen Tätigkeit als Interviewer behält. Bestehende Skalen sollten hierzu kritisch geprüft und gegebenenfalls adaptiert werden. Außerdem sollte die Erfassung von Kontrollüberzeugungen vor dem Feldbeginn erfolgen und der bereits angesprochene Einfluss

von Selektionseffekten ist zu berücksichtigen, da häufig eine hohe Fluktuation in den Interviewerstämmen zu beobachten ist, die stark an die Arbeitsbedingungen gekoppelt scheint und kausale Aussagen erschwert (vgl. Link 2006). Auch sollte die Interviewererfahrung aus anderen Instituten und Unternehmen berücksichtigt werden, da längere Beschäftigungen bei einem einzigen Arbeitgeber eher die Ausnahme darstellen und außerdem Erfahrungen aus verschiedenen Kontexten womöglich auch unterschiedliche Effekte aufweisen können (vgl. Hüfken/Schäfer 2003). Und schließlich sollten zukünftig unbedingt wiederkehrende Befragungen der Interviewer durchgeführt werden, da nur so kausale Einflüsse und sich wandelnde Zusammenhänge zweifelsfrei aufgedeckt und beschrieben werden können. Die von uns vorgestellten Querschnittsanalysen sind hierzu nicht in der Lage. Somit sollte die widerlegte Hypothese einer größeren Bedeutung von Kontrollüberzeugungen unter erfahrenen Interviewern (Hypothese 3) nicht als zwingendes Indiz für einen statischen Zusammenhang von Kontrollüberzeugung und Interviewerfolg gewertet werden. Längsschnittsanalysen können hier sicherlich interessante Verläufe aufdecken. Dies gilt ebenso für den komplexen Zusammenhang von Kontrollüberzeugungen und Interviewerfolg wie für alle weiteren Veränderungen im Laufe der in Analysen bislang vernachlässigten Interviewerkarrieren.

Literatur

- ADM, 2006: Jahresbericht 2006. Frankfurt a. M.: Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. http://www.adm-ev.de/pdf/Jahresbericht_06.pdf (01.09.2008).
- Andrisani, P. J. und G. Nestel, 1976: Internal-External Control as Contributor to and Outcome of Work Experience. *Journal of Applied Psychology* 61: 156-165.
- Blohm, M., J. Hox und A. Koch, 2006: The Influence of Interviewers' Contact Behavior on the Contact and Cooperation Rate in Face-to-Face Household Surveys. *International Journal of Public Opinion Research* 19: 97-111.
- Couper, M. P. und R. M. Groves, 1992: The Role of Interviewer in Survey Participation. *Survey Methodology* 18: 263-277.
- Curtin, R., S. Presser und E. Singer, 2005: Changes in Telephone Survey Nonresponse over the Past Quarter Century. *Public Opinion Quarterly* 69: 87-98.
- de Leeuw, E., J. Hox, G. Snijders und W. de Heer, 1998: Interviewer Opinions, Attitudes and Strategies Regarding Survey Participation and their Effect on Response. S. 239-248 in: A. Koch und R. Porst (Hg.): *Nonresponse in Survey Research*. Mannheim: GESIS.
- Furnham, A. und H. Steele, 1993: Measuring Locus of Control: A Critique of General, Children's, Health- and Work-related Locus of Control Questionnaires. *British Journal of Psychology* 84: 443-479.
- Gabler, S. und S. Häder, 2002: *Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland*. Münster: Waxmann Verlag.
- Hansen, K. M., 2006: The Effects of Incentives, Interview Length, and Interviewer Characteristics on Response Rates in a CATI-Study. *Journal of Public Opinion Research* 19: 112-121.
- Hox, J. und E. de Leeuw, 2002: The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison. S. 103-120 in: R. M. Groves, D. A. Dillman, J. L. Eltinge und R. J. A. Little (Hg.): *Survey Nonresponse*. New York: Wiley.

- Hüfken, V. und A. Schäfer, 2003: Zum Einfluss stimmlicher Merkmale und Überzeugungsstrategien der Interviewer auf die Teilnahme in Telefonumfragen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 55: 321-339.
- Jakoby, N. und R. Jacob, 1999: Messung von internen und externen Kontrollüberzeugungen in allgemeinen Bevölkerungsumfragen. *ZUMA-Nachrichten* 45: 61-71.
- Judge, T. A. und J. E. Bono, 2001: Relationship of Core Self-Evaluations Traits - Self-Esteem, Generalized Self-Efficacy, Locus of Control, and Emotional Stability - With Job Satisfaction and Job Performance: A Meta-Analysis. *Journal of Applied Psychology* 86: 80-92.
- Kohler, M., 2005: Stichprobenkonzeption und Stichprobenrealisierung. S. 31-57 in: R. Koch-Institut (Hg.): *Erster telefonischer Gesundheitssurvey des Robert Koch-Instituts - Methodische Beiträge*. Berlin: Eigenverlag.
- Kohler, M., A. Rieck und S. Borch, 2005: Methode und Design des telefonischen Gesundheitssurveys 2003. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 48: 1224-1229.
- Link, M. W., 2006: Predicting the Persistence and Performance of Newly Recruited Telephone Interviewers. *Field Methods* 18: 305-320.
- Muhonen, T. und E. Torkelson, 2004: Work Locus of Control and its Relationship to Health and Job Satisfaction from a Gender Perspective. *Stress and Health* 20: 21-28.
- Oksenberg, L., L. Coleman und C. F. Cannell, 1986: Interviewers' Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly* 50: 97-111.
- Olson, K. und A. Peytchev, 2007: Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes. *Public Opinion Quarterly* 71: 273-286.
- Rasbach, J., F. Steele, W. Browne und B. Prosser, 2005: *A User's Guide to MLwiN. Version 2.0*. Bristol: University of Bristol.
- Rotter, J. B., 1966: Generalized Expectancies for Internal versus External Control of Reinforcement. *Psychological Monographs* 80: 1-28.
- Schmich, P., 2006: *Interviewereffekte bei telefonischen Umfragen*. FU Berlin: Diplomarbeit.
- Singer, E., M. R. Frankel und M. B. Glassman, 1983: The Effect of Interviewer Characteristics and Expectations on Response. *Public Opinion Quarterly* 47: 68-83.
- Siu, O.-I., P. E. Spector, C. L. Cooper und I. Donald, 2001: Age Differences in Coping and Locus of Control: A Study of Managerial Stress in Hong Kong. *Psychological and Aging* 16: 707-710.
- Spector, P. E., 1982: Behavior in Organizations as a Function of Employee's Locus of Control. *Psychological Bulletin* 91: 482-497.
- Stephens, A. und J. Wardle, 2001: Locus of Control and Health Behaviour Revisited: A Multivariate Analysis of Young Adults from 18 Countries. *British Journal of Psychology* 92: 659-672.
- Twisk, J. W. R., 2006: *Applied Multilevel Analysis*. Cambridge: Cambridge University Press.

Korrespondenzadressen: Patrick Schmich, Sabine Schiller-Born
Robert Koch-Institut
Postfach 65 02 61
D-13302 Berlin
SchmichP@rki.de
Schiller-BornS@rki.de

Christian Weidmann
Mannheimer Institut für Public Health
Medizinische Fakultät Mannheim
Universität Heidelberg
Ludolf-Krehl-Str. 7-11
D-68167 Mannheim
christian.weidmann@medma.uni-heidelberg.de

Zufallsbefragung von Ausländern auf Basis des Ausländer- zentralregisters

*Erfahrungen bei der
Repräsentativbefragung
„Ausgewählte Migranten-
gruppen in Deutschland
2006/2007“ (RAM)*

Random Sampling Based on the Central Register for Foreigners

*Experiences from the
Representative Survey
"Selected Groups of
Migrants in Germany
2006/2007"*

*Christian Babka von Gostomski und
Monika Pupeter*

Zusammenfassung

In Deutschland haben rund 15 Millionen Menschen einen Migrationshintergrund. Das Ausländerzentralregister ist eine Datenquelle, die einen Teil der Personen mit Migrationshintergrund – die Ausländer – umfasst, sodass es für diese als Grundlage der Stichprobenziehung dienen kann. Dies wurde erstmals mit der Repräsentativbefragung „Ausgewählte Migranten­gruppen in Deutschland 2006/2007“ realisiert. Diese Befragung führte das Bundesamt für Migration und Flüchtlinge, Nürnberg, in Zusammenarbeit mit TNS Infratest Sozialforschung, München, durch. Insgesamt wurden 1.544 Türken, 972 Personen aus dem ehemaligen Jugoslawien, 746 Italiener, 677 Griechen und 637 Polen befragt. Auch wenn sich die Stichprobenziehung aus dem Ausländerzentralregister und die Adressenbeschaffung über Ausländerbehörden über einen Zeitraum von einem halben Jahr hinzog, zahlte sich dieser lange Vorlauf aus. Denn Ausländer konnten zielgenauer als bei den Voruntersuchungen mittels einer zufallsbasierten Stichprobe befragt werden. Die Abweichungen zwischen realisierter Auswahl und Grundgesamtheit

Abstract

A population of approximately 15 million people with migration background is currently living in Germany. In order to be able to describe this population of migrants and foreigners, samples drawn on a random basis are the preferable approach. The Central Register for Foreigners holds a database which, as yet, has not been used for drawing a partial sample of persons with a migration background. For the first time a sample was drawn from the Central Register for Foreigners for the representative survey "Selected Groups of Migrants in Germany 2006/2007", operated by the Federal Office for Migration and Refugees, Nuremberg, in co-operation with TNS Infratest Sozialforschung, Munich. Within the framework of this survey, 1,544 Turks, 972 people from former Yugoslavia, 746 Italians, 677 Greeks and 637 people from Poland were interviewed. Even though the drawing from this database required a good deal of time and work in advance, it shows that the additional work is worth the effort, since it offers a more detailed focus than similar studies for

hinsichtlich der Merkmale Geschlecht, Alter und Bundesland sind nur gering. Die Ausschöpfung liegt im Rahmen dessen, was bei vergleichbaren Einwohnermeldeamtstichproben erzielt wird. Im Rahmen der Durchführung der Studie wurde zudem deutlich, dass bei Ausländerbefragungen Übersetzungen des Frageprogramms unabdingbar sind.

reaching the targeted group of foreigners in Germany. The divergence between realized selection and the sample of the population based on the attributes sex, age and Federal Länder can be considered as minor. Furthermore, it becomes apparent that for interviews with foreigners, translated questionnaires are essential.

1 Einleitung

Mit Vorlage des Nationalen Integrationsplans, den Bekenntnissen zu Zuwanderung und Integration als Teil der Geschichte Deutschlands und dem Erfordernis einer nachhaltigen Integrationspolitik (Die Bundesregierung 2007: 12) ist das Thema der Zuwanderung und Integration weiter in den Fokus der Öffentlichkeit gelangt. Die Erkenntnisse um Zuwanderung und Integration wurden vielfach durch große Studien zu Migranten geprägt.¹ Das Ergebnis aus dem Mikrozensus 2005, dass rund 15 Millionen Personen mit Migrationshintergrund in Deutschland leben (Statistisches Bundesamt 2007b), macht deutlich, dass die Forschungstätigkeiten rund um die immer heterogener werdenden Migrantengruppen in Zukunft eher zu- als abnehmen werden.

Bei der Vertiefung der Erkenntnisse über die in Deutschland lebenden Migrantengruppen ist die empirische Sozialforschung gefordert, die jeweiligen Zielgruppen möglichst ‚repräsentativ‘ zu erfassen und zu beschreiben. Inzwischen werden eine Reihe von Verfahren verwendet, um Personen mit Migrationshintergrund zu identifizieren und Stichproben zu ziehen. Als Beispiele seien toponomastische oder onomastische Verfahren (etwa Humbert/Schneiderheinze 2000; Salentin 2007), Einwohnermeldeamtstichproben (Ette et al. 2007) und die Kombination von mindestens zwei dieser Verfahren (etwa Diehl 2005: 18f.; Granato 1999: 58; Santacreu Fernandez/Rother/Braun 2006) genannt (allgemein auch: Salentin 1999).

Gemeinhin wird es aber bei der Ziehung einer Zufallsauswahl aus einer Grundgesamtheit als ideal angesehen, wenn von der Grundgesamtheit eine vollständige Liste vorliegt (Schnell/Hill/Esser 2005: 271). Für einen Teil der Personen

1 Es wird aus Gründen der besseren Lesbarkeit im Folgenden die männliche Form verwendet.

mit Migrationshintergrund in Deutschland, nämlich die in Deutschland lebenden Ausländer, liegt mit dem Ausländerzentralregister (AZR) eine solche Liste vor. Bei der Stichprobenziehung zur Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“ (RAM 2006/2007) wurde erstmals eine große Befragung mit rund 4.500 Befragten auf der Basis einer Zufallsauswahl aus dem AZR realisiert. Die Prozesse, Möglichkeiten und Restriktionen, die mit dieser Art der Stichprobenziehung verbunden sind, werden im Folgenden dargestellt.

2 Zielsetzung der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“ (RAM 2006/2007)

RAM 2006/2007 steht in loser Tradition mit den vom damaligen Bundesministerium für Arbeit und Sozialordnung in Auftrag gegebenen und in den Jahren 1980 (Mehrländer et al. 1981), 1985 (König et al. 1986), 1995 (Mehrländer/Ascheberg/Ueltzhöffer 1996) und 2001 (Venema/Grimm 2002a, 2002b) durchgeführten Befragungen, die sich mit den jeweils größten Gruppen der in den alten Bundesländern lebenden Ausländer befassten. Alle vier Vorgängeruntersuchungen zu RAM 2006/2007 kamen mit Hilfe von Quotenauswahlen in bestimmten Gebieten in den alten Bundesländern und West-Berlin zustande (König/Schultze/Wessel 1986: 7ff.; Mehrländer et al. 1981: 8ff.; Mehrländer/Ascheberg/Ueltzhöffer 1996: 12ff.; Venema/Grimm 2002a: 12ff.). Befragungsmethode waren jeweils persönlich-mündliche Interviews mit Hilfe von Papierfragebögen.

Dem Bundesamt für Migration und Flüchtlinge (BAMF) in Nürnberg wurde vom Bundesministerium des Innern die Verantwortlichkeit für eine Neuauflage der Studie übertragen. Das BAMF betreibt auch – neben vielen anderen Aufgaben – wissenschaftliche Begleitforschung über Migrationsfragen zur Gewinnung analytischer Aussagen für die Steuerung der Zuwanderung. Das BAMF beauftragte im Juni 2006 TNS Infratest Sozialforschung (Infratest) in München, um in enger Zusammenarbeit die Studie durchzuführen.

Gegenüber den Vorgängeruntersuchungen wurden bei RAM 2006/2007 Änderungen der methodischen Anlage vorgenommen. Erstens wurde eine Zufallsstichprobe aus dem AZR gezogen, bei der auch Ausländer aus den neuen Bundesländern berücksichtigt wurden. Zweitens fand die Untersuchung als Computer-unterstützte persönliche Befragung (Computer Assisted Personal Interviewing, CAPI) statt.

RAM 2006/2007 soll

- den mit der Ausländer-, Einwanderungs- und Integrationspolitik befassten Stellen umfassende und aktuelle Daten zu den kulturellen und wirtschaftlichen Lebensverhältnissen der fünf größten Ausländergruppen in Deutschland zur Verfügung stellen,
- Informationen bereitstellen, welche die amtliche Statistik nicht liefern kann, und
- vertiefende Analysen zu Integrations- und Migrationsprozessen bei heute in Deutschland ansässigen Ausländern ermöglichen (zu Ergebnissen: Babka von Gostomski 2008; Bundestag Drucksache 16/10155 2008: 7f.; Haug 2008; Worbs 2008).

Mit den Daten zu RAM 2006/2007 sind Analysen zu folgenden Themenbereichen möglich: Soziodemographische Merkmale der fünf Ausländergruppen; Bildung; Berufsausbildung, berufliche Situation und Einkommen; Sprache; Wohnsituation; Haushalt und familiäre Situation; soziale Kontakte; Bindung an Deutschland und das Heimatland; Religion; Werte, persönliche Vorstellungen und Ansichten; Probleme in Deutschland, Beratungsanspruchnahme, -wünsche.

3 Das Ausländerzentralregister (AZR)

3.1 Grundinformationen über das AZR

Eine zentrale Rolle für die Stichprobenziehung und die Bestimmung der Zielpersonen für RAM 2006/2007 spielte das AZR (ausführlich zum AZR: Bundesamt für Migration und Flüchtlinge 2008a). Das Bundesverwaltungsamt in Köln verarbeitet und nutzt die AZR-Daten im Auftrag des BAMF. Das AZR ist eine bundesweite, personenbezogene Datei mit Daten von allen Personen, die nicht Deutsche im Sinne des Art. 116 Abs. 1 GG sind, das heißt, nicht die deutsche Staatsangehörigkeit besitzen. Dazu zählen auch Staatenlose und Personen mit ungeklärter Staatsangehörigkeit. Deutsche, die zugleich eine fremde Staatsangehörigkeit besitzen, zählen – entsprechend dieser staatsrechtlichen Definition – nicht zur ausländischen Bevölkerung (zu Personen wie Stationierungstreitkräften oder Diplomaten, die nicht über das AZR erfasst werden: Statistisches Bundesamt 2007a: 6). Die Daten von Personen etwa mit einem deutschen und gleichzeitig einem türkischen Pass sind nicht im AZR gespeichert. Die Zulieferung der Daten erfolgt unter anderem durch über 600 Ausländerbehörden (zur Verteilung der Ausländerbehörden und deren Bezirke: Bundesamt für Migration und Flüchtlinge 2007: 35).

Im Rahmen der Bemühungen zur Verbesserung der Datenqualität erfolgte in den Jahren 2000 bis 2004 eine Bereinigung des AZR (Bundesministerium des Innern/Bundesamt für Migration und Flüchtlinge 2005: 103). Der Unterschied von 7.334.765 Ausländern in Deutschland zum Jahresende 2003 zu 6.717.115 Ausländern zum Jahresende 2004 ist zum Großteil auf diese Bereinigung zurückzuführen (Statistisches Bundesamt 2007a: 4).

Bei einer Datei wie dem AZR, die ständig mit neuen Daten von sehr vielen Ausländerbehörden und anderen Stellen beliefert wird, ist es fast unvermeidlich, dass Unplausibilitäten auftauchen (zu problematischen Sachverhalten im AZR: Statistisches Bundesamt 2007a: 6). So wurden etwa im Zuge der Vorarbeiten zur Bestimmung der Grundgesamtheit für RAM 2006/2007 einige Dubletten, also Personen, die zweimal im AZR abgespeichert wurden, gestrichen.

Im Datenbestand des AZR werden die Daten von Ausländern erfasst, die sich nicht nur vorübergehend in Deutschland aufhalten. Daher bietet das AZR eine ideale Basis für die Stichprobenziehung dieser Ausländer. Die Anschrift des Ausländers ist jedoch kein Speichersachverhalt im AZR. Daher ist für eine Befragung die Beteiligung der Ausländerbehörden erforderlich.

Nur dem BAMF ist es in Ausnahmefällen gestattet, im Rahmen seiner gesetzlichen Aufgaben zur Begleitforschung der Zuwanderungsprozesse von und nach Deutschland das AZR zur Stichprobenziehung zu nutzen.² Externen Forscherteams außerhalb des BAMF steht diese Möglichkeit der Stichprobenziehung über das AZR nicht offen.

3.2 Restriktionen bei einer Stichprobenziehung über das AZR

Die Vorgängeruntersuchungen von 1980 bis 2001 hatten das Ziel, Personen mit einer ausländischen Staatsangehörigkeit zu befragen (zur Befragung 2001 diesbezüglich: Venema/Grimm 2002a: 12). Als eine Referenzstatistik, um die Quotenvorgaben nach Gebietseinheiten, Geschlecht, Alter, Familienstand und Aufenthaltsdauer zur Auswahlbestimmung zu ermitteln, wurde das AZR herangezogen (zur Befragung 1995 diesbezüglich: Mehrländer/Ascheberg/Ueltzhöffer 1996: 13f.; zur Befragung 2001 diesbezüglich: Venema/Grimm 2002a: 13). Mit dem Aufbau einer

2 Die Stichprobenziehung zu RAM 2006/2007 erfolgte dabei im Rahmen einer so genannten ‚Vorgriffsregelung‘, das heißt, dass das BAMF im Vorgriff auf eine zukünftig zu erwartende Regelung im AZR-Gesetz (Wissenschaftsklausel) Personendaten aus dem AZR unter bestimmten Voraussetzungen abfragen durfte. Ein Entwurf für eine Wissenschaftsklausel wurde durch das BAMF formuliert; diese wird voraussichtlich in das AZR-Gesetz aufgenommen. Damit würde auch weiterhin nur dem BAMF, nicht aber externen Forschungsgruppen, der Zugang zu Personendaten aus dem AZR unter bestimmten Voraussetzungen möglich.

Forschungsgruppe im BAMF war abzusehen, dass erstmals das AZR direkt – und nicht nachträglich wie noch 1995 und 2001 über eine Anpassung an Randverteilungen – Verwendung finden konnte (Mehrländer/Ascheberg/Ueltzhöffer 1996: 13f.; Venema/Grimm 2002a: 13ff.). Deswegen wurde von einer Quotenauswahl für RAM 2006/2007 Abstand genommen.

Mit einer Stichprobenziehung aus dem AZR sind Restriktionen verbunden, die klar benannt werden müssen, um die Ergebnisse aus RAM 2006/2007 richtig zu interpretieren. Denn, wie oben schon dargestellt, es sind nur Daten von Ausländern im staatsrechtlichen Sinn im AZR. Der Begriff ‚Migrant‘, der in der Studie verwendet wird, stellt damit auf die Staatsangehörigkeit zu einem bestimmten Stichtag und nicht auf den Migrationshintergrund ab. Verallgemeinerungen auf alle Personen mit Migrationshintergrund in Deutschland verbieten sich damit. Vereinfacht ausgedrückt ist aus neueren Studien zu Personen mit Migrationshintergrund bekannt, dass unter den Personen mit Migrationshintergrund die eingebürgerten Personen etwa in den Integrationsbereichen Bildung, Berufsausbildung, Erwerbstätigkeit, berufliche Stellung, Einkommen und Sprachkenntnisse in der Regel eine bessere Integrationsbilanz vorzuweisen haben als Nicht-Deutsche (zu Ergebnissen diesbezüglich etwa: Brenke 2008: 503f.; Bundesministerium für Arbeit und Soziales 2008: 142ff.; Ministerium für Generationen, Familie, Frauen und Integration des Landes Nordrhein-Westfalen 2008: 140ff.; Salentin/Wilkening 2003; Seifert 2007; Statistisches Bundesamt 2007b, 2008). Da man nun aber um diese Unterschiede im Stand der Integration zwischen Eingebürgerten und Ausländern ohne die deutsche Staatsangehörigkeit weiß, ist es gerechtfertigt, sich den bezüglich einiger Integrationsaspekte weniger gut integrierten Ausländern in einem eigenen großen Survey zu widmen.

Eine weitere Restriktion bezüglich der Vergleichbarkeit ergibt sich aus der Verwendung des AZR zur Gewichtung bei der Vorgängeruntersuchung 2001 bei gleichzeitiger Beibehaltung der Personen im Datensatz, die schon eine deutsche und vielleicht auch gleichzeitig eine ausländische Staatsangehörigkeit hatten (Venema/Grimm 2002a: 19 und 69).³ Trendanalysen sind damit zwischen 2001 und RAM 2006/2007 nicht möglich. Zudem wurden bei der Befragungsmethode Änderungen vorgenommen (von persönlich-mündlichen Interviews mit Papierfragebogen 2001 zu Computer-unterstützter persönlicher Befragung 2006/2007). Im Zuge der Ände-

3 Ob bei deutscher Staatsangehörigkeit die bisherige Staatsangehörigkeit aufgegeben oder behalten wurde, wurde 2001 nicht erfragt (Venema/Grimm 2002a: 69). Je nach Ausländergruppe gaben zwischen 9% und 15% der Befragten eine deutsche Staatsangehörigkeit an und sind damit keine Ausländer im staatsrechtlichen Sinne, wie es dem AZR zugrunde liegt (siehe Kapitel 3.1).

rung der Befragungsmethode, aber auch im Zuge der Abstimmung im Prozess der Instrumentenentwicklung, wurden zudem zahlreiche Fragen und Antwortkategorien geändert. Die in Kapitel 2 angesprochene ‚lose Tradition‘ zu den Vorgängeruntersuchungen bezieht sich daher auf bestimmte Themenbereiche, die tangiert, nicht aber identisch zu 2001 repliziert wurden.

4 Durchführung von RAM 2006/2007 mit Hilfe des AZR

4.1 Grundgesamtheit

Die Elemente der Grundgesamtheit für RAM 2006/2007 sind eine Auswahl aus der Gesamtheit aller am 30. Juni 2006 im AZR erfassten Ausländer. Die Auswahlkriterien bezogen auf den Stichtag 30. Juni 2006 waren dabei folgende:

- türkische, griechische, italienische oder polnische Staatsangehörigkeit oder eine Staatsangehörigkeit eines Nachfolgestaates des ehemaligen Jugoslawien (Bosnien-Herzegowina, Kroatien, Slowenien, Serbien und Montenegro, Jugoslawien oder Mazedonien),
- Vollendung des 14. Lebensjahrs und jünger als 80 Jahre,
- Mindestaufenthaltsdauer von zwölf Monaten in Deutschland und
- gesicherter Aufenthaltsstatus.⁴

Die Tabelle 1 stellt die Grundgesamtheit für RAM 2006/2007 vergleichend für die Merkmale Geschlecht, gruppiertes Alter zum Stichtag 30. Juni 2006 und Bundesland, differenziert nach den fünf Gruppen (Staatsangehörigkeiten) dar.

4 Hinter der Formulierung ‚gesicherter Aufenthaltsstatus‘ zum Stichtag 30. Juni 2006 verbergen sich verschiedene Sachverhalte der Kategorien ‚Aufenthaltslaubnis‘, ‚EU-Aufenthaltsrecht‘, ‚Niederlassungserlaubnis‘ sowie alte Titel nach Ausländergesetz (zu einer genauen Auflistung auch Babka von Gostomski 2007: 56f.).

Tabelle 1 Grundgesamtheit zum Stichtag 30. Juni 2006, Anzahl der Fälle (Basis) und Spaltenprozent

	Ausländergruppen					Gesamt
	Türken	Personen aus dem ehemaligen Jugoslawien	Italiener	Griechen	Polen	
Basis	1.369.810	721.216	451.476	257.527	241.442	3.041.471
Männer	53,3	51,7	60,1	54,6	42,7	53,2
Frauen	46,7	48,3	39,9	45,4	57,3	46,8
15 bis 19 Jahre	10,8	6,2	5,6	4,8	3,5	7,8
20 bis 24 Jahre	8,8	7,7	7,6	7,0	8,0	8,1
25 bis 29 Jahre	12,5	11,8	9,2	9,5	16,0	11,9
30 bis 34 Jahre	13,9	13,4	11,2	12,1	18,0	13,5
35 bis 39 Jahre	11,4	10,3	11,8	12,2	13,1	11,4
40 bis 44 Jahre	10,4	7,1	10,9	10,1	10,9	9,7
45 bis 49 Jahre	5,4	5,7	10,6	9,2	11,6	7,1
50 bis 54 Jahre	4,9	8,5	9,5	7,4	9,6	7,0
55 bis 59 Jahre	6,2	13,0	8,6	8,2	5,3	8,3
60 bis 64 Jahre	7,0	7,7	5,8	7,1	1,8	6,6
65 bis 69 Jahre	5,4	5,2	4,8	6,6	1,0	5,0
70 bis 74 Jahre	2,5	2,6	2,8	4,1	0,7	2,6
75 bis 79 Jahre	0,8	0,9	1,4	1,7	0,4	1,0
Schleswig-Holstein	2,0	1,1	0,7	1,2	3,1	1,6
Hamburg	3,2	2,9	1,0	2,1	5,3	2,9
Niedersachsen	6,1	4,6	4,5	5,2	9,9	5,7
Bremen	1,6	0,7	0,3	0,4	1,5	1,1
Nordrhein-Westfalen	33,9	21,8	24,2	31,0	27,3	28,8
Hessen	10,3	10,7	11,9	9,7	10,1	10,6
Rheinland-Pfalz	4,0	3,7	5,1	2,4	5,0	4,0
Baden-Württemberg	17,2	26,1	30,9	24,0	8,9	21,3
Bayern	13,4	21,6	14,9	19,4	13,5	16,1
Saarland	0,7	0,6	3,4	0,3	0,8	1,1
Berlin	7,0	5,2	2,2	3,0	9,1	5,7
Brandenburg	0,1	0,2	0,1	0,2	1,9	0,3
Mecklenburg-Vorpommern	0,1	0,1	0,1	0,2	0,7	0,1
Sachsen	0,2	0,3	0,3	0,5	1,6	0,4
Sachsen-Anhalt	0,1	0,2	0,1	0,3	0,6	0,2
Thüringen	0,1	0,1	0,2	0,2	0,5	0,1

Quelle: Sonderauszählung aus dem AZR: Alle türkischen, griechischen, italienischen und polnischen Personen, die am Stichtag 30. Juni 2006 das 14. Lebensjahr vollendet und jünger als 80 Jahre waren, eine Mindestaufenthaltsdauer von zwölf Monaten in Deutschland und einen gesicherten Aufenthaltsstatus hatten. Gleiches gilt für Bürger eines Nachfolgestaates des ehemaligen Jugoslawien (Bosnien-Herzegowina, Kroatien, Slowenien, Serbien und Montenegro, Jugoslawien oder Mazedonien).

4.2 Angestrebte Brutto-Stichprobe und angestrebte Interviews

In Absprache des BAMF mit Infratest fiel die Wahl auf ein zweistufiges Verfahren, um zu einer Zufallsstichprobe zu gelangen (ausführlich: Pupeter 2007). In der ersten Stufe wurden dabei Ausländerbehörden ausgewählt, aus denen in einem zweiten Schritt die Ziehung der Zielpersonen erfolgte. Des Weiteren wurde ein disproportionaler Ansatz gewählt. Von den Türken und von Personen aus dem ehemaligen Jugoslawien sollten anteilmäßig etwas weniger Personen, als prozentual in Deutschland vertreten sind, befragt werden und von den anderen drei Ausländergruppen etwas mehr. Hintergrund dieser Zielsetzung war, auch für die ‚kleineren‘ Ausländergruppen eine hinreichend große Fallzahl zu realisieren, damit multivariate Auswertungen sowie Auswertungen von Untergruppen möglich sind.

In der ersten Stufe wurde, für die fünf Ausländergruppen getrennt, eine geschichtete (siehe dazu unten) Zufallsauswahl von Ausländerbehörden gezogen. Einige der über 600 Ausländerbehörden in Deutschland fanden keine Berücksichtigung, da sie als zentrale Aufnahmestellen Asylsuchende führen und daher für die hier interessierenden Ausländergruppen nicht relevant waren. Insgesamt gingen in die Zufallsauswahl 593 Ausländerbehörden ein, wobei je nach Ausländergruppe unterschiedlich viele Ausländerbehörden Berücksichtigung fanden.

In diesem Prozess war auch die Größenordnung der letztlich zu befragenden Personen der fünf Ausländergruppen zu beachten. Es war das Ziel, rund 1.500 Türken, 1.000 Personen aus dem ehemaligen Jugoslawien, 750 Italiener, 650 Griechen und 600 Polen zu befragen. In Tabelle 2 sind diese in der Spalte ‚angestrebte Interviews, Anzahl‘ zu finden.

Tabelle 2 Grundgesamtheit, angestrebte Brutto-Stichprobe, angestrebte und realisierte Interviewzahl

	Grundgesamtheit		Angestrebte Brutto-Stichprobe		Angestrebte Interviews		Realisierte Interviews	
	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent
Türken	1.369.810	45,0	10.472	33,3	1.500	33,3	1.544	33,8
Personen aus dem ehemaligen Jugoslawien	721.216	23,7	7.000	22,3	1.000	22,2	972	21,2
Italiener	451.476	14,8	5.208	16,6	750	16,7	746	16,3
Griechen	257.527	8,5	4.536	14,4	650	14,4	677	14,8
Polen	241.442	7,9	4.200	13,4	600	13,3	637	13,9
Gesamt	3.041.471	100	31.416	100	4.500	100	4.576	100

Die Ziehung der Stichprobe wurde über Sample-Points organisiert. Die Größe eines Sample-Points bemaß sich daran, wie viele Nettointerviews pro Point sinnvollerweise durchgeführt werden sollten. Grundsätzlich sollten, zwecks Homogenisierung der Varianz innerhalb des Points, alle Sample-Points gleich groß sein. Aus der Zahl der angestrebten Nettointerviews pro Sample-Point ergab sich die Zahl der Sample-Points je Ausländergruppe. Pro Point wurden acht Nettointerviews angesetzt. Damit ergab sich folgende Anzahl der Sample-Points je Ausländergruppe:

Türken	$1.500 / 8 = 187$	(Sample-Points)
Personen aus dem ehemaligen Jugoslawien	$1.000 / 8 = 125$	„
Italiener	$750 / 8 = 93$	„
Griechen	$650 / 8 = 81$	„
Polen	$600 / 8 = 75$	„

Die Schichtung der Ausländerbehörden erfolgte nach Bundesland und Größenklassen. Zur Bestimmung der Größenklassen lag mittels Sonderzählungen des AZR auf Ebene der Ausländerbehörden die Anzahl der Zielpersonen unterteilt nach Geschlecht vor. Ausländerbehörden, in denen weniger als 250 Personen der betreffenden Ausländergruppen gemeldet waren, wurden aus Datenschutzgründen von der Stichprobenziehung ausgeschlossen.⁵ Die Anzahl der Sample-Points wurde zufällig auf die geschichteten Ausländerbehörden verteilt. Durch Allokation ergab sich eine sinnvolle Größe von 190 einzubeziehenden Ausländerbehörden.

Mit der Rechengröße ‚acht Nettointerviews pro Sample-Point‘ pro Ausländergruppe kommt man bei siebenfacher Übersetzung auf 56 zu ziehende Adressen pro Point. Für die Übersetzung wurde ein so hoher Ansatz gewählt, da mit der Art der Stichprobenziehung auf Basis des AZR bisher noch keine Erfahrungen bezüglich der Qualität der Adressen und der Kooperationsbereitschaft der Ausländerbehörden bei der nachgelagerten Adressanspielung vorlagen. Um einen reibungslosen Ablauf planen zu können, wurden mehr Adressen angefordert als später eingesetzt werden sollten. Insgesamt ergaben sich mit 561 Sample-Points 31.416 potentiell nutzbare Adressen (siehe Tabelle 2 zur Übersicht und detailliert Tabelle 3).

5 Diese Einschränkung hat zur Konsequenz, dass Ausländer mit einem Wohnsitz in einem Ausländerbehördenbezirk mit einem nur sehr geringen Anteil ihrer Staatsangehörigkeit in der Stichprobe nicht vorkommen. Allerdings ist auch bei Einwohnermeldeamtsstichproben mit ähnlichen Restriktionen zu rechnen. So konnten beim von Infratest im Auftrag des Bundesinstituts für Bevölkerungsforschung durchgeführten Generations and Gender Survey (GGG) aus datenschutzrechtlichen Gründen nur Gemeinden mit mindestens 200 türkischen Zielpersonen berücksichtigt werden (Ette et al. 2007: 13).

Tabelle 3 Angestrebte Brutto-Stichprobe, Anzahl der Fälle (Basis) und Spaltenprozentage

	Ausländergruppen					Gesamt
	Türken	Personen aus dem ehemaligen Jugoslawien	Italiener	Griechen	Polen	
Basis	10.472	7.000	5.208	4.536	4.200	31.416
Männer	53,0	50,5	59,8	52,5	43,8	52,3
Frauen	47,0	49,5	40,2	47,5	56,2	47,7
15 bis 19 Jahre	10,5	5,7	5,5	5,3	4,0	7,0
20 bis 24 Jahre	9,0	8,4	8,0	7,7	9,6	8,6
25 bis 29 Jahre	13,1	11,9	9,4	9,6	17,2	12,2
30 bis 34 Jahre	13,5	12,7	11,3	11,2	18,4	13,3
35 bis 39 Jahre	11,3	9,6	11,6	11,7	12,7	11,2
40 bis 44 Jahre	11,0	7,1	10,5	9,6	10,7	9,8
45 bis 49 Jahre	5,5	5,6	9,9	8,8	9,8	7,3
50 bis 54 Jahre	4,6	8,5	10,6	7,7	9,7	7,6
55 bis 59 Jahre	6,3	13,9	8,7	8,3	4,8	8,5
60 bis 64 Jahre	6,7	7,6	5,3	7,7	1,5	6,1
65 bis 69 Jahre	5,1	5,2	4,7	6,3	0,6	4,6
70 bis 74 Jahre	2,7	2,8	3,1	4,2	0,7	2,7
75 bis 79 Jahre	0,8	1,0	1,3	1,8	0,4	1,0
Schleswig-Holstein	2,1	1,6		1,2	4,0	1,8
Hamburg	3,2	3,2	1,1	2,5	5,3	3,0
Niedersachsen	5,9	4,0	4,3	6,2	9,3	5,7
Bremen	1,6	0,8		1,2	1,3	1,1
Nordrhein-Westfalen	33,7	21,6	24,7	30,9	26,7	28,2
Hessen	10,7	11,2	11,8	9,9	10,7	10,9
Rheinland-Pfalz	4,3	3,2	5,4	2,5	5,3	4,1
Baden-Württemberg	17,1	26,4	31,2	23,5	8,0	21,2
Bayern	13,4	21,6	15,1	18,5	13,3	16,2
Saarland	0,5	0,8	3,2		1,3	1,1
Berlin	7,0	4,8	2,2	2,5	9,3	5,3
Brandenburg					1,3	0,2
Mecklenburg-Vorpommern					1,3	0,2
Sachsen	0,5	0,8	1,1	1,2	2,7	1,1

Quelle: Sonderauszählung aus dem AZR: Nach Zufall aus 190 Ausländerbehörden gezogene türkische, griechische, italienische und polnische Personen, die am Stichtag 30. Juni 2006 das 14. Lebensjahr vollendet und jünger als 80 Jahre waren, eine Mindestaufenthaltsdauer von zwölf Monaten in Deutschland und einen gesicherten Aufenthaltsstatus hatten. Gleiches gilt für Bürger eines Nachfolgestaates des ehemaligen Jugoslawien (Bosnien-Herzegowina, Kroatien, Slowenien, Serbien und Montenegro, Jugoslawien oder Mazedonien).

Als Ergebnis der ersten Stufe der Stichprobenziehung lag damit eine Auswahl der Ausländerbehörden sowie die Zahl der zu ziehenden Bruttoadressen je Ausländerbehörde in Abhängigkeit von der Zahl der Points pro Ausländerbehörde vor.

Beim Vergleich der Grundgesamtheit (siehe Tabelle 1) mit der angestrebten Brutto-Stichprobe (siehe Tabelle 3) ergibt sich insgesamt eine hinreichend proportionale Abbildung der Grundgesamtheit. Werden zwei Prozentpunkte und mehr als größere Abweichung interpretiert, dann sind lediglich griechische Männer in der angestrebten Brutto-Stichprobe unterrepräsentiert. Man erkennt zudem, dass die Bundesländer Sachsen-Anhalt und Thüringen aufgrund ihres geringen Anteils von Ausländern der fünf Gruppen in diesen Bundesländern beim Auswahlprozess nicht berücksichtigt wurden.⁶

Die Liste der im ersten Schritt ausgewählten 190 Ausländerbehörden mit der jeweiligen Zahl der zu ziehenden Bruttoadressen je Ausländergruppe wurde dann zwecks Durchführung der zweiten Stufe der Stichprobenziehung im BAMF bearbeitet.⁷ Dabei wurden die Personen nach den Vorgaben zufällig aus dem AZR gezogen. Im August 2006 schrieb das BAMF die 190 Ausländerbehörden an, in denen die in Tabelle 3 dargestellten 31.416 Personen zu finden sein sollten. Jede Ausländerbehörde erhielt eine Liste mit dem Nachnamen, dem Vornamen, dem Geburtsdatum, dem Geschlecht, der Staatsangehörigkeit, der Nummer der Ausländerbehörde, dem Kurznamen der Ausländerbehörde sowie, wenn vorhanden, der individuellen Ausländerbehördennummer der Person. Es war nun an den Ausländerbehörden, die Listen mit der Adresse der Befragten zu vervollständigen und diese dann an das BAMF zurückzusenden.

Insgesamt lieferten in der vorgesehenen Zeitspanne zwischen August 2006 und Anfang November 2006 183 Ausländerbehörden ihre vervollständigten Listen an das BAMF. Einige wenige Ausländerbehörden sahen sich für diesen Zeitraum außerstande, die Listen zu vervollständigen oder die Listen trafen zu spät ein.⁸ Die Überprüfung und Bearbeitung der zurückgesandten Listen und Angaben im

6 Für die neuen Bundesländer wurden bei der angestrebten Brutto-Stichprobe die Türken, die Personen aus dem ehemaligen Jugoslawien, die Italiener und die Griechen ausschließlich aus Sachsen rekrutiert, während die Polen (5,3% aller Polen) aus drei der fünf neuen Bundesländer stammen (siehe Tabelle 3).

7 Die Anzahl der berücksichtigten Ausländerbehörden variierte damit nach den Befragtengruppen: Für die Türken wurde bei der angestrebten Brutto-Stichprobe auf Angaben von 136 Ausländerbehörden, für die Personen aus dem ehemaligen Jugoslawien von 96, die Italiener von 82, die Griechen von 64 und die Polen auf Angaben von 60 Ausländerbehörden im AZR zurückgegriffen.

8 Auch bei bundesweiten Befragungen, die auf Einwohnermeldeamtsstichproben beruhen, gelingt es in der Regel nicht, von allen ursprünglich gezogenen Einwohnermeldeämtern Daten zu den Zielpersonen zu erhalten. Beispielsweise fanden beim ALLBUS 2006 148 Gemeinden Berücksichtigung. Sechs Gemeinden mussten jedoch „ausgetauscht werden, da die örtlichen Behörden nicht anfechtbare Absagegründe mitteilten“ (Wasmer/Scholz/Blohm 2007: 56).

BAMF benötigte Zeit, um letztlich Infratest eine leicht zu bearbeitende Liste für die Durchführung der Interviews zur Verfügung zu stellen. Hierbei muss bedacht werden, dass die Ausländerbehörden relativ frei darin sind, wie sie und auf welchem IT-System sie ihre internen Aufgaben organisieren. Zudem variieren die Systeme zuweilen von Bundesland zu Bundesland. Es wurden daher zum Teil sehr umfangreiche Datenkonvertierungs- und Angleichungsarbeiten durchgeführt, um die endgültige Liste zu erstellen.

Mitte November 2006 erfolgte der Versand der angeglichenen Liste von 27.436 Personen an Infratest. Diese enthielt den Nachnamen, den Vornamen, das Geschlecht, die Staatsangehörigkeit, die Nummer der Ausländerbehörde, den Kurznamen der Ausländerbehörde sowie die Adresse der jeweiligen Zielperson mit dem Straßennamen und der Hausnummer sowie der Postleitzahl und dem Ortsnamen. Nicht enthalten war das Geburtsdatum. Bei der Zahl ist zu beachten, dass erstens Angaben von einigen Ausländerbehörden fehlten, die auf Wunsch des jeweiligen Datenschutzbeauftragten direkt an den Datenschutzbeauftragten von Infratest lieferten. Die Differenz zu den 31.416 Personen der angestrebten Brutto-Stichprobe ergibt sich zweitens aus dem oben angesprochenen Ausfall von einigen Ausländerbehörden. Drittens konnten einige gelieferte Daten zu Einzelpersonen nicht verwendet werden.⁹ Viertens waren bei den an das BAMF zurückgesandten Listen aus den 183 Ausländerbehörden einige dabei, die weniger Personen enthielten als angefragt.

Infratest führte bei der aus der BAMF-Liste einerseits und aus den von einigen Ausländerbehörden direkt an Infratest gesandten Listen andererseits zusammengespülten Liste mit insgesamt 29.116 Adressen weitere Prüfungen auf Vollständigkeit und Dubletten durch, wonach die Adressen von 28.537 Personen als verwendbar erachtet wurden. Diese Adressen verteilen sich folgendermaßen auf die fünf Ausländergruppen: 9.613 Türken, 6.440 Personen aus dem ehemaligen Jugoslawien, 4.736 Italiener, 4.090 Griechen und 3.658 Polen. Von den ursprünglich anvisierten 31.416 Personen der angestrebten Brutto-Stichprobe wurden also letztlich die Adressen und Angaben zu 28.537 Personen als verwendbar erachtet. Dies sind 90,8%. Rund ein Zehntel der vorgesehenen Adressen erwies sich damit als fehlerhaft oder konnte aufgrund der Nichtzusendung oder einer verspäteten Zusendung durch einige Ausländerbehörden nicht genutzt werden (siehe hierzu auch noch zusammenfassend Tabelle 5 in Kapitel 4.3).

9 Bei bundesweiten Einwohnermeldeamtsstichproben müssen in der Regel ebenfalls die von den Behörden gelieferten Listen bereinigt werden, etwa um doppelte Adressen oder um solche Adressen, die offensichtlich keine Privathaushalte beinhalten (Wasmer/Scholz/Blohm 2007: 57).

4.3 Ausschöpfung und realisierte Interviews

Zur Realisierung der Befragung zog Infratest aus den 28.537 Adressen eine Stichprobe von 11.220 Personen, die am 7. Dezember 2006 angeschrieben wurden. Da sich von den angeschriebenen Adressen wieder einige Adressen als nicht verwendbar erwiesen, erfolgte Ende Januar 2007 und Mitte März 2007 der Versand weiterer Anschreiben an Zielpersonen und die Weitergabe weiterer Adressen an Interviewer.

In Tabelle 4 (Zeile 1: Sample-Points) findet sich zunächst noch einmal die bereits im Kapitel 4.2 erklärte Anzahl der Sample-Points wieder.¹⁰ Zunächst wurden 20 Personen pro Sample-Point angeschrieben (siehe Tabelle 4, Zeile 2). Bei 1.065 Personen stellten sich die Anschreiben als nicht zustellbar heraus, das entspricht 9,5% aller Adressen (siehe Tabelle 4, Zeile 4). Besonders hoch ist dabei der Anteil der Nichterreichbaren bei polnischen Personen (20,5%). Bei Personen, bei denen die Post nicht zustellbar war, wurde keine Realisierungschance eines Interviews gesehen, diese werden als ‚qualitätsneutrale Ausfälle‘ (QNA) bezeichnet.¹¹ Um die angestrebte Fallzahl nicht zu gefährden, wurde eine Nachbesetzung mit Reserveadressen aus dem gleichen Point, welche per Zufallsauswahl gezogen wurden, vorgenommen (‚1. QNA Ersatz‘, siehe Tabelle 4, Zeile 5). So ergeben sich wiederum 11.220 Adressen, die als ‚Bruttoadressen für Feldeinsatz‘ bezeichnet werden (siehe Tabelle 4, Zeile 6).¹²

Insgesamt erwiesen sich in der Feldphase 1.650 (= 14,7%) der 11.220 Adressen als falsch, die Zielperson war an dieser Adresse unbekannt oder unbekannt verzogen, gehörte nicht zur Zielgruppe oder war verstorben (siehe im Einzelnen: Tabelle 4, Zeilen 7 bis 10). Diese Ausfälle sind als ‚stichprobenneutrale Ausfälle‘ (siehe Tabelle 4, Zeile 11) zusammengefasst. Auch hier ist bei den Polen festzustellen, dass die entsprechenden Adressen sich zu 15,2% als falsch herausstellten. Bei den anderen vier Gruppen erwiesen sich hingegen nur 3,1% bis 7,2% der Bruttoadressen für den Feldeinsatz als falsch.

Auch diese 1.650 stichprobenneutralen Ausfälle wurden im Laufe der Feldphase fast gänzlich durch weitere 1.627 Adressen aus den verbleibenden 16.252 Adressen ersetzt (siehe Tabelle 4, Zeile 12: ‚2. QNA Ersatz‘). Insgesamt wurden mit 13.912 Adressen nur knapp die Hälfte der bereinigten Brutto-Stichprobe von 28.537 Adressen für die Untersuchung genutzt (siehe Tabelle 4, Zeile 13).

10 In Tabelle 4 steht ‚QNA‘ für ‚qualitätsneutrale Ausfälle‘, ‚ZP‘ für ‚Zielperson‘.

11 Der Ausfallgrund ‚qualitätsneutraler Ausfall (QNA)‘ beziehungsweise ‚stichprobenneutraler Ausfall‘ ist umstritten. Eine Person, die verzogen ist, gehört trotzdem weiterhin zur Stichprobe, jedoch ohne Chance auf Realisierung. Aus pragmatischen Gründen und da ‚qualitätsneutrale Ausfälle‘ beziehungsweise ‚stichprobenneutrale Ausfälle‘ gängige Begriffe sind, werden sie auch hier verwendet.

12 Das erste Interview der Hauptuntersuchung fand am 13. Dezember 2006, das letzte am 30. April 2007 statt. Insgesamt kamen 293 Interviewer zum Einsatz, die im Durchschnitt 16 Interviews durchführten. Die Interviews dauerten im Durchschnitt 44 Minuten (Pupeter 2007: 9ff.).

Tabelle 4 Ausschöpfungübersicht

	Türken		Personen aus dem ehe- möglichen Jugoslawien		Italiener		Griechen		Polen		Gesamt	
	n	%	n	%	n	%	n	%	n	%	n	%
1) Sample-Points	187		125		93		81		75		561	
2) Angeschriebene Adressen je Point	20		20		20		20		20		20	
3) Ursprünglich verschickte Anschreiben	3.740	100	2.500	100	1.860	100	1.620	100	1.620	100	11.220	100
4) Anschreiben nicht zustellbar	256	6,8	195	7,8	160	8,6	147	9,1	147	20,5	1.065	9,5
5) 1. QNA Ersatz	256	6,8	195	7,8	160	8,6	147	9,1	147	20,5	1.065	9,5
6) Bruttoadressen für Feldeinsatz	3.740	100	2.500	100	1.860	100	1.620	100	1.620	100	11.220	100
7) Adresse falsch	260	7,0	181	7,2	58	3,1	87	5,4	87	15,2	814	7,3
8) ZP unbekannt/unbekannt verzogen	236	6,3	170	6,8	134	7,2	131	8,1	131	8,5	798	7,1
9) ZP gehört nicht zur Zielgruppe	10	0,3	5	0,2	2	0,3	4	0,2	4	0,0	21	0,2
10) ZP verstorben	5	0,1	4	0,2	5	0,3	3	0,2	3	0,0	17	0,2
11) Stichprobenneutrale Ausfälle	511	13,7	360	14,4	199	10,7	225	13,9	225	23,7	1.650	14,7
12) 2. QNA Ersatz	508	13,6	355	14,2	192	10,3	222	13,7	222	23,3	1.627	14,5
13) Insgesamt benötigte Adressen	4.504	120,4	3.050	122,0	2.212	118,9	1.989	122,8	1.989	143,8	13.912	124,0
14) Tatsächliches Stichprobenbrutto	3.737	100	2.495	100	1.853	100	1.617	100	1.617	100	11.197	100
15) Im Haushalt niemanden angetroffen	478	12,8	312	12,5	226	12,2	183	11,3	183	16,0	1.438	12,8
16) ZP nicht angetroffen	190	5,1	135	5,4	93	5,0	85	5,3	85	8,0	623	5,6
17) ZP verreist/im Urlaub	156	4,2	57	2,3	72	3,9	78	4,8	78	2,1	394	3,5
18) ZP krank/nicht in der Lage	76	2,0	65	2,6	35	1,9	33	2,0	33	2,0	239	2,1
19) ZP hat keine Zeit	328	8,8	274	11,0	172	9,3	143	8,8	143	6,4	1.013	9,0
20) ZP nicht bereit, sonstiger Grund	785	21,0	610	24,4	456	24,6	355	22,0	355	19,2	2.493	22,3
21) Sprachschwierigkeiten	156	4,2	42	1,7	22	1,2	45	2,8	45	1,3	284	2,5
22) Nicht endgültig bearbeitet	18	0,5	27	1,1	31	1,7	18	1,1	18	2,4	130	1,2
23) Ausfälle gesamt	2.187	58,5	1.522	61,0	1.107	59,7	940	58,1	940	57,4	6.614	59,1
24) Durchgeführte Interviews	1.550	41,5	973	39,0	746	40,3	677	41,9	677	42,6	4.583	40,9
25) Nicht auswertbar	6	0,2	1	0,0	0	-	0	-	0	-	7	0,0
26) Realisierte Interviews	1.544	41,3	972	39,0	746	40,3	677	41,9	677	42,6	4.576	40,9

Als ‚tatsächliches Stichprobenbrutto‘ (siehe Tabelle 4, Zeile 14) werden jene Personen der insgesamt benötigten Adressen (siehe Tabelle 4, Zeile 13) bezeichnet, bei denen das Anschreiben zustellbar war und die nicht zu den stichprobenneutralen Ausfällen zu zählen sind: Das sind 11.197 Personen.¹³ Betrachtet man die Gründe der Ausfälle (Zeilen 15 bis 22) beim tatsächlichen Stichprobenbrutto, dann stellt man auch hier wieder fest, dass die Polen häufiger bei den Rubriken ‚Im Haushalt niemanden angetroffen‘ (16% des ‚tatsächlichen Stichprobenbruttos‘ gegenüber 11,3% bei den Griechen, Zeile 15 in Tabelle 4) und ‚Zielperson nicht angetroffen‘ (8% bei den Polen gegenüber 5% bei den Italienern) zu finden sind, als es bei den anderen vier Gruppen der Fall ist. Bei Personen aus dem ehemaligen Jugoslawien finden sich hingegen verstärkt die Ausfallgründe ‚Zielperson hat keine Zeit‘ (11% gegenüber 6,4% bei den Polen) sowie ‚Zielperson nicht bereit, sonstiger Grund‘ (24,4% gegenüber 19,2% bei den Polen). Auch bei den Italienern fällt ein vergleichsweise hoher Wert beim letztgenannten Ausfallgrund auf (24,6%). In der Feldarbeit zeigte sich zudem, dass polnische Männer häufiger nicht (mehr) unter der angegebenen Adresse zu finden waren. Dies deutet darauf hin, dass sie entweder nicht in dem Umfang in Deutschland leben, wie sie im AZR ausgewiesen werden, oder in der Zeitspanne von der Stichprobenziehung bis zum Interviewversuch schon wieder unter einer anderen Adresse lebten.

Obwohl der Fragebogen nicht nur in Deutsch vorlag, sondern auch übersetzt wurde (siehe auch Kapitel 4.5), ergibt sich bei einem weiteren Ausfallgrund, den ‚Sprachschwierigkeiten‘ (siehe Tabelle 4, Zeile 21), ein leicht erhöhter Anteil von Ausfällen bei türkischen Zielpersonen: 4,2% des tatsächlichen Stichprobenbruttos kamen hier wegen Sprachschwierigkeiten nicht zustande, bei Italienern und Polen betragen diese Ausfälle hingegen nur 1,2% und 1,3%.

Von den durchgeführten Interviews erwiesen sich schließlich nur wenige, nämlich sieben, als nicht verwertbar: Es wurde dabei ermittelt, dass fälschlicherweise ein Familienangehöriger der Zielperson befragt wurde.

Insgesamt ist hinsichtlich der verschiedenen Stufen der Ausschöpfung festzustellen, dass Interviews mit Polen zu einem vergleichsweise hohen Anteil nicht zustande kamen, da schon die Adressen sich als nicht auffindbar oder falsch herausstellten. Diese fehlerhaften oder nicht mehr aktuellen Adressen zeigten sich dann auch bei jenen Adressen, die sich nicht schon im Vorfeld, bei den sogenannten qualitätsneutralen oder stichprobenneutralen Ausfällen, als nicht verwendbar erwiesen. Auch bei jenen Adressen, mit deren Hilfe versucht wurde, ein Interview in die Wege

13 Die Zahlen bei ‚tatsächliches Stichprobenbrutto‘ errechnen sich aus Zeile 13 minus der Summe aus Zeile 5 und 11.

zu leiten, zeigten sich größere Anteile bei den Polen, bei denen überhaupt niemand oder nicht die gesuchte Person im Haushalt zu finden war. Da in die Ausschöpfung als Basis nur das tatsächliche Stichprobennetto einfließt, ergibt sich bei den Polen eine vergleichsweise gute Ausschöpfung von 42,6%, da bei ihnen die Ausfallgründe ‚verreist/im Urlaub‘, ‚krank/nicht in der Lage‘, ‚keine Zeit‘, ‚nicht bereit, sonstiger Grund‘ und ‚Sprachschwierigkeiten‘ vergleichsweise selten von den Interviewern angegeben wurden. Die von allen Gruppen geringste Ausschöpfung von 39% ergab sich bei den Personen aus dem ehemaligen Jugoslawien, da diese verstärkt keine Zeit hatten, nicht bereit waren, an dem Interview teilzunehmen, oder sich bei ihnen sonstige Gründe für das Nichtzustandekommen eines Interviews einstellten. Darüber erklärt sich auch, dass die Zielvorgabe von 1.000 Interviews bei Personen aus dem ehemaligen Jugoslawien (siehe Tabelle 2 bei angestrebten Interviews) mit 972 realisierten Interviews in dem vorgegebenen Zeitrahmen nicht erreicht wurde. Bei den Italienern liegt die Ausschöpfung von 40,3% mit 746 realisierten Interviews nur geringfügig unter der anvisierten Größe von 750. Bei den anderen drei Gruppen wurden hingegen die Zielvorgaben mit 1.544 realisierten Interviews mit Türken, 677 mit Griechen und 637 mit Polen deutlich übertroffen. Mit einer Ausschöpfung von 40,9% liegt RAM 2006/2007 im Rahmen dessen, was heutzutage üblicherweise bei anspruchsvollen Befragungen auf Basis einer Zufallsstichprobe erzielt wird.¹⁴

Aus den bisher dargestellten Schritten lässt sich ein Fazit zur Abschätzung der Qualität der über das beschriebene AZR-Verfahren gewonnenen Adressen und der Möglichkeit der Realisierung von Interviews bei Zielpersonen aus dem AZR ziehen. Bis ein Interviewer an der Haustür einer Zielperson klingeln kann, sind drei Hürden zu überwinden. Erstens müssen von den Ausländerbehörden mit der Anschrift ergänzte Adressen vorliegen, die als verwendbar eingeschätzt werden. In 90,8% der Fälle wurde dies erreicht (siehe Tabelle 5, Zeile 3). Zweitens sind die Befragten vorab über Anschreiben über die bevorstehende Befragung zu informieren. Dies gelang bei 90,5% der Fälle (siehe Tabelle 5, Zeile 6). Und drittens muss sich schließlich die Adresse beim Versuch der Befragung als der Zielperson zugehörig und richtig erweisen, was bei 85,3% der Fälle eintrat (siehe Tabelle 5, Zeile 9). Werden die Adressen, die sich bis hierhin als für den Feldeinsatz geeignet erweisen, insgesamt

14 So wurde beim ALLBUS 2006 in Westdeutschland eine Ausschöpfung von 40,2% und in Ostdeutschland von 42,8% erzielt (Wasmer/Scholz/Blohm 2007: 66). Ette et al. (2007: 16) berichten für den GGS für die Gruppe der türkischen Migranten von einer Ausschöpfung von 34,1%. Bei einer älteren Analyse von Koch (1998: 66) ergaben sich allerdings keine empirischen Belege dafür, dass Stichprobenverzerrungen vom Umfragen mit höherer Ausschöpfung geringer sind als bei solchen mit geringer Ausschöpfung (zu einer ähnlichen Einschätzung: Schneekloth/Leven 2003, zu einer kritischen Haltung bezüglich geringer Ausschöpfungsquoten aber: Mohler/Koch/Gabler 2003).

quantifiziert, dann sind letztlich 70,1% als verwendbar anzusehen.¹⁵ Auch wenn das Klingelschild oder die Briefkastenbeschriftung die Zielperson als prinzipiell erreichbar ausweist, ist danach aber noch ein weiterer Schritt zum letztlich realisierten Interview nötig. Berücksichtigt man die Ausfälle, die beim Interviewversuch noch entstehen können, dann kommt man nach den Erkenntnissen aus RAM 2006/2007 bei 28,7% der laut AZR potenziellen Zielpersonen zu einem realisierten Interview (siehe Tabelle 5, Zeile 14).

Vergleichbare, bundesweite Studien mit Ausländern sind in Deutschland rar. Beim Generations and Gender Survey (GGG), der 2006 mittels einer Einwohnermeldeamtsstichprobe von in Deutschland gemeldeten türkischen Staatsangehörigen im Alter zwischen 18 und 79 Jahren zustande kam, lässt sich der Anteil der verwendbaren Adressen mit 83,2% beziffern (Ette et al. 2007: 17).¹⁶ Die Ausschöpfung lag bei 34,1%. Dabei war der Anteil der potenziellen Zielpersonen des verbleibenden Stichprobenbruttos, bei denen im Haushalt niemand angetroffen wurde, mit 26,3% deutlich höher als bei Türken in RAM 2006/2007 (12,8%). Insgesamt wurde beim GGS mit 28,3% der laut Einwohnermeldebehörden potenziellen Zielpersonen mit türkischer Staatsangehörigkeit ein verwertbares Interview geführt. Betrachtet man zudem den ALLBUS 2006, dessen Grundgesamtheit aus Deutschen und Ausländern ab 18 Jahren besteht und ebenfalls mittels einer Einwohnermeldeamtsstichprobe durchgeführt wurde (Wasmer/Scholz/Blohm 2007: 53ff.), dann kommt man im Westen auf 89% (Ost: 90,8%) verwendbare Adressen, eine Ausschöpfung von 40,2% (Ost: 42,8%), und damit konnte bei 35,8% (Ost: 38,9%) der laut Einwohnermeldebehörden potenziellen Zielpersonen ein auswertbares Interview geführt werden. Auch wenn solche Vergleiche immer aufgrund verschiedenster Gründe – zu denken ist etwa an unterschiedlich lange Feldzeiten, Befragungsdauer, Themen der Befragungen, unterschiedliche Darstellung der Ausschöpfung (hierzu detailliert: Ette et al. 2007: 17; Wasmer/Scholz/Blohm 2007: 68) – nur sehr eingeschränkt möglich sind, so zeigt sich doch, dass auch bei bundesweiten Einwohnermeldeamtsstichproben mit zwischen 10% und 20% nicht verwertbaren Adressen zu rechnen ist.¹⁷

15 Die 70,1% errechnen sich über $((\text{Zeile 2}/\text{Zeile 1}) * (\text{Zeile 5}/\text{Zeile 4}) * (\text{Zeile 8}/\text{Zeile 7}))$ in der Spalte ‚Gesamt‘ (siehe Tabelle 5, Zeile 10).

16 Dies ist unter anderem darauf zurückzuführen, dass beim GGS keine Vorab-Verifikation der Adresse mittels des Versands des Ankündigungsschreibens stattgefunden hat.

17 Dabei können auch scheinbare Nebensächlichkeiten eine Rolle spielen: Granato (1999: 58) etwa stellte bei einer auf Mannheim begrenzten Studie fest, dass aus einer Einwohnermeldeamtsstichprobe gezogene Nachnamen von türkischen Zielpersonen zu 78% am Klingelschild der Wohnung zu finden waren, bei Deutschen waren es hingegen 93%.

Tabelle 5 Abschätzung über die Qualität des AZR-Gesamtverfahrens zur Adressenbeschaffung und zur Realisierung von Interviews

	Ausländergruppen					Gesamt
	Türken	Personen aus dem ehemaligen Jugoslawien	Italiener	Griechen	Polen	
1) Angestrebte Brutto-Stichprobe	10.472	7.000	5.208	4.536	4.200	31.416
2) Angekommene und verwendbare Adressen	9.613	6.440	4.736	4.090	3.658	28.537
3) % der angekommenen und verwendbaren Adressen	91,8	92,0	90,9	90,2	87,1	90,8
4) Ursprünglich verschickte Anschreiben	3.740	2.500	1.860	1.620	1.620	11.220
5) Anschreiben zugestellt	3.484	2.305	1.700	1.473	1.473	10.155
6) % der zugestellten Anschreiben	93,2	92,2	91,4	90,9	90,9	90,5
7) Bruttoadressen für Feldeinsatz	3.740	2.500	1.860	1.620	1.620	11.220
8) Stichprobenbrutto	3.229	2.140	1.661	1.395	1.395	9.570
9) % des Stichprobenbruttos	86,3	85,6	89,3	86,1	86,1	85,3
10) Abschätzung des Prozentsatzes der verwendbaren Adressen für den Feldeinsatz	73,8	72,6	74,2	70,6	68,2	70,1
11) Tatsächliches Stichprobenbrutto	3.737	2.495	1.853	1.617	1.495	11.197
12) Realisierte Interviews	1.544	972	746	677	637	4.576
13) % der realisierten Interviews	41,3	39,0	40,3	41,9	42,6	40,9
14) Abschätzung des Prozentsatzes der realisierbaren Interviews	30,5	28,3	29,9	29,6	29,1	28,7

Sowohl bundesweite Einwohnermeldeamtsstichproben als auch das Verfahren über das AZR haben mit dem Problem zu kämpfen, an die laut jeweiligem Register dort anzutreffenden Personen zu gelangen. Ein Teil der Zieladressen erweist sich als nicht verwertbar. Bei RAM 2006/2007 erwies sich dieses Problem als gravierender im Vergleich mit dem GGS. Bei als zutreffende Zieladressen erachteten Adressen kann es dann vorkommen, dass im Haushalt niemand angetroffen wird. Dies kam häufiger beim GGS als bei RAM 2006/2007 vor.¹⁸

18 Auf die Möglichkeit, Stichproben für Ausländer- beziehungsweise Migrantenbefragungen aus Telefonbüchern zu ziehen, kann an dieser Stelle nicht im Detail eingegangen werden (zu Möglichkeiten und Problemen bei Zuwandererstichproben aus dem Telefonbuch generell: Salentin 2002; zur Vorschaltung onomastischen Verfahren: Humbert/Schneiderheinze 2002: 188ff.; Salentin 2007: 37f.; zu Übersichten bezüglich Ausfällen von der telefonischen Ausgangsstichprobe bis zu einem realisierten Interview etwa: Häder 2000: 10 sowie mehrere Beiträge in Gabler/Häder 2002).

4.4 Abweichungen zwischen Grundgesamtheit und realisierten Interviews

Tabelle 6 zeigt die letztlich realisierten Interviews bei RAM 2006/2007 nach den Merkmalen Geschlecht, Altersgruppen zum Stichtag 30. Juni 2006 und Bundesland ähnlich wie bei Tabelle 1.

Werden Beträge der Prozentpunkte, die sich aus der Differenz der Tabellen 1 und 6 errechnen, mit einem Wert größer als 2,0 als bemerkenswerte Abweichungen interpretiert, dann fallen einige Sachverhalte auf. Entsprechend dem gehäuften Vorkommen von falschen Adressen oder nicht mehr zutreffenden Adressen insbesondere bei polnischen Männern ist es nicht verwunderlich (siehe Kapitel 4.3), dass polnische Frauen bei RAM 2006/2007 gegenüber der Verteilung in der Grundgesamtheit überrepräsentiert sind.¹⁹ Bezüglich der Bundesländer ist markant, dass insbesondere Personen aus Baden-Württemberg bei den realisierten Interviews unterrepräsentiert sind. Da sich diese Verteilung bei allen fünf Ausländergruppen zeigt, ist dafür wohl eine geringere Intervieweffizienz in einer der Regionen Baden-Württembergs verantwortlich.

Überproportional häufig wurden bei RAM 2006/2007 Türken, Italiener und Griechen aus Nordrhein-Westfalen befragt. Befragte Polen aus Niedersachsen sind hingegen überrepräsentiert und solche aus Hessen unterrepräsentiert. Zudem sind bei RAM 2006/2007 Türken, Personen aus dem ehemaligen Jugoslawien und Italiener aus Bayern stärker vertreten als in der Grundgesamtheit. Bezüglich des Alters lässt sich sagen, dass 15- bis 19-Jährige etwas stärker bei der Studie als in der Grundgesamtheit repräsentiert sind.

Insgesamt zeigt sich aber, dass sich nur wenige ausgeprägte Abweichungen zwischen Grundgesamtheit und realisierten Interviews bei RAM 2006/2007 einstellen. In der Regel gelang damit die Abbildung der Geschlechter – Ausnahme Polen –, der Altersgruppen und auch nach den Bundesländern – Ausnahmen: Baden-Württemberg, Nordrhein-Westfalen und Bayern – bei allen fünf Ausländergruppen gut.²⁰

19 Anhand der realisierten Interviews bei RAM 2006/2007 lässt sich zeigen, dass unter polnischen Frauen viele sind, die im Rahmen der Familienzusammenführung als Ehepartnerin nach Deutschland kamen. Polnische Männer geben hingegen häufiger an, dass die Arbeitssuche ein Grund für die Zuwanderung nach Deutschland gewesen sei. Zudem sind erwerbstätige polnische Männer vergleichsweise oft nur befristet angestellt (Babka von Gostomski 2008: 23; Haug 2008: 27). Es liegt daher nahe anzunehmen, dass die polnischen Heiratsmigrantinnen ortsfester sind als polnische Männer mit wechselnden Beschäftigungen und damit deren Adresse schneller veraltet.

20 Da die Grundgesamtheit mit Hilfe der Auswertungen aus dem AZR bekannt war, ließen sich bei RAM 2006/2007 anhand der Schichten ‚Geschlecht‘, ‚Staatsangehörigkeit zum Stichtag 30. Juni 2006‘, ‚Bundesland‘, ‚Alter zum Stichtag 30. Juni 2006‘ Gewichtungsfaktoren erstellen, die die realisierten Interviews an die Strukturen der Grundgesamtheit anpassen.

Tabelle 6 Realisierte Interviews bei RAM 2006/2007, Anzahl der Fälle (Basis) und Spaltenprozent

	Ausländergruppen					Gesamt
	Türken	Personen aus dem ehemaligen Jugoslawien	Italiener	Griechen	Polen	
Basis	1.544	972	746	677	637	4.576
Männer	52,2	51,3	60,5	54,4	34,4	51,2
Frauen	47,8	48,7	39,5	45,6	65,6	48,8
15 bis 19 Jahre	13,9	8,2	6,8	8,4	5,5	9,6
20 bis 24 Jahre	9,0	9,3	6,2	8,6	7,4	8,3
25 bis 29 Jahre	12,4	11,5	7,0	9,2	14,6	11,1
30 bis 34 Jahre	13,5	12,3	9,8	8,9	18,1	12,6
35 bis 39 Jahre	12,2	9,9	13,5	12,3	13,8	12,2
40 bis 44 Jahre	12,5	9,7	11,5	11,2	12,6	11,6
45 bis 49 Jahre	5,5	7,3	9,0	10,5	10,7	7,9
50 bis 54 Jahre	4,8	7,9	9,4	8,0	8,6	7,2
55 bis 59 Jahre	5,2	13,2	9,2	7,7	5,0	7,9
60 bis 64 Jahre	5,7	5,7	6,2	6,1	1,6	5,2
65 bis 69 Jahre	3,6	3,0	6,4	5,8	1,1	3,9
70 bis 74 Jahre	1,6	1,2	3,1	3,0	0,6	1,8
75 bis 79 Jahre	0,1	0,8	1,9	0,6	0,5	0,7
Schleswig-Holstein	1,4	0,7		1,0	3,3	1,2
Hamburg	3,9	3,1	1,2	1,6	5,7	3,2
Niedersachsen	5,4	5,9	7,4	5,8	15,7	7,3
Bremen	2,1	0,6		0,7	0,9	1,1
Nordrhein-Westfalen	38,9	20,2	29,0	35,6	29,2	31,4
Hessen	10,7	14,4	10,1	11,4	5,3	10,7
Rheinland-Pfalz	2,8	3,0	4,4	2,2	5,2	3,4
Baden-Württemberg	12,2	19,4	25,6	17,6	4,4	15,6
Bayern	15,8	25,5	17,6	20,4	11,8	18,3
Saarland	0,5	0,2	2,5		1,9	0,9
Berlin	6,2	6,0	0,8	1,5	10,4	5,1
Brandenburg					2,5	0,3
Mecklenburg-Vorpommern					1,3	0,2
Sachsen	0,2	1,0	1,5	2,2	2,5	1,2

Quelle: RAM 2006/2007 (ungewichtet): Türkische, griechische, italienische und polnische Personen, die am Stichtag 30. Juni 2006 das 14. Lebensjahr vollendet und jünger als 80 Jahre waren, eine Mindestaufenthaltsdauer von zwölf Monaten in Deutschland und einen gesicherten Aufenthaltsstatus hatten. Gleiches gilt für Bürger eines Nachfolgestaates des ehemaligen Jugoslawien (Bosnien-Herzegowina, Kroatien, Slowenien, Serbien und Montenegro, Jugoslawien oder Mazedonien).

Im Zeitraum zwischen Stichprobenziehung (Stichtag 30. Juni 2006) und Befragung (letztes Interview am 30. April 2007) kann es zu Einbürgerungen gekommen sein. Bei der Gruppe derjenigen Befragten, die im AZR als Türken, als Personen aus dem ehemaligen Jugoslawien oder als Polen erfasst waren, findet sich ein nennenswerter Anteil von Personen, die zum Befragungszeitpunkt (auch) die deutsche Staatsangehörigkeit hatten (Türken: 3,6%, ehemalige Jugoslawen: 3,6%, Polen: 3,8%), während es unter Italienern (1,5%) und Griechen (1,2%) nur sehr wenige sind (Prozentanteile jeweils ungewichtet).

4.5 Zur Notwendigkeit fremdsprachiger Fragebögen

In vielen nationalen Befragungen, wie dem ALLBUS 2006 oder dem Mikrozensus 2005, werden, obwohl auch Ausländer befragt werden, die Befragten nur mittels deutschsprachiger Befragungsinstrumente interviewt (zum ALLBUS 2008 diesbezüglich: Wasmer/Scholz/Blohm 2007: 68; zum Mikrozensus 2005 diesbezüglich: GESIS 2008, dort Gliederungspunkt ‚Datenerhebung‘; zu weiteren methodischen Problemen wie der Untererfassung von Aussiedlern oder zu Proxyinterviews beim Mikrozensus etwa: Seifert 2008: 11f.; Zühlke 2008).²¹ Anhand RAM 2006/2007 lässt sich abschätzen, wie hoch der Anteil der Ausländer der fünf berücksichtigten Gruppen ist, die nur eingeschränkt oder gar nicht allein mit einem deutschen Befragungsinstrument hätten befragt werden können. Denn das Frageprogramm wurde in die Hauptsprachen der fünf Ausländergruppen übersetzt. Um die Sprachkenntnisse der Interviewten nicht nur durch die Selbstbewertung der Befragten abschätzen zu können, wurden zusätzliche Angaben erhoben, die der Interviewer zur Einschätzung des Sprechens der deutschen Sprache des Befragten und zu der Nutzung der Übersetzungshilfe beziehungsweise der Hinzuziehung eines Übersetzers machte. Werden alle drei Sachverhalte zusammen betrachtet, dann nahmen die türkischen Befragten recht häufig (25,7%) die Übersetzungshilfe in Anspruch und bei ihnen wurde am häufigsten (24%) ein Übersetzer zur Hilfe geholt.²² Auch in der Einschätzung der Interviewer weisen die türkischen Befragten die geringsten deutschen Sprachkenntnisse auf: 38,9% von ihnen haben aus der Sicht der Interviewer nur mittelmäßige bis sehr schlechte deutsche Sprachkenntnisse. Aber

21 Beim ALLBUS 2006 sprachen 2,1% der Zielpersonen der bereinigten Stichprobe nicht ausreichend deutsch, sodass kein Interview mit ihnen geführt werden konnte (Wasmer/Scholz/Blohm 2007: 68). Die Ausländer im ALLBUS 2006 sind damit bezüglich des Integrationsindikators ‚deutsche Sprachkenntnisse‘ besser einzuschätzen als alle Ausländer in Deutschland.

22 Beim GGS wurden 33% der Interviews mit türkischen Migranten unter Einsatz der Übersetzungen durchgeführt (Ette et al. 2007: 15).

auch bei den anderen vier Befragungsgruppen kam die Übersetzungshilfe bei etwa einem Fünftel zum Einsatz, ein Übersetzer wurde bei etwa jedem zehnten Interview benötigt. Insgesamt macht dies darauf aufmerksam, dass es bei allen fünf Ausländergruppen bei Befragungen nötig ist – auch wenn bei der Stichprobenziehung darauf Wert gelegt wurde, dass die Befragten eine Mindestaufenthaltsdauer von zwölf Monaten in Deutschland aufwiesen (siehe Kapitel 4.1) –, dass eine Übersetzung des Frageprogramms vorliegt.

5 Zusammenfassung und Ausblick

In der empirischen Sozialforschung herrscht das Postulat, dass Zufallsauswahlen die zu präferierenden Auswahlverfahren sind (Deutsche Forschungsgemeinschaft 1999: 47; Schumann 2000: 84). Bei den Vorgängerstudien zu RAM 2006/2007 wurden jeweils Quotenauswahlen zugrunde gelegt. Insofern wurde bei RAM 2006/2007 von der Anlage her – über die Stichprobenziehung aus dem AZR – die Grundlage zu einer höheren Datenqualität und zu einer besseren Basis zur Verallgemeinerung der Ergebnisse auf die befragten Ausländergruppen in Deutschland im Vergleich zu den Vorgängeruntersuchungen gelegt. Es zeigte sich dann auch, dass sich mit dieser Vorgehensweise bei den Zielgruppen eine gute Anpassung zwischen Stichprobe und Grundgesamtheit erzielen lässt.

Allerdings nimmt eine Stichprobenziehung mit Hilfe des AZR einen nicht geringen Zeitraum in Anspruch. Bei der ersten Realisierung einer Stichprobe aus dem AZR im Rahmen von RAM 2006/2007 gingen der konkreten Stichprobenziehung im Juli 2006 langwierige Gespräche und Abstimmungsprozesse, etwa über Datenschutzaspekte und die Bereitstellung eines Individualdatensatzes, mit dem das Zufallsverfahren durchgeführt werden konnte, voraus (zu Vorüberlegungen zu RAM 2006/2007 schon: Haug/Worbs/Theuer 2005). Nach der Klärung dieser Fragen dauerte es weitere rund fünf Monate, bis über die Ausländerbehörden die Listen der ausgesuchten Zielpersonen erstellt, vereinheitlicht und dem Befragungsinstitut zur Realisierung der Interviews zur Verfügung gestellt werden konnten. Die Schritte, die dabei zu bearbeiten waren, seien hier nur noch einmal stichpunktartig aufgeführt:

- Bearbeitung des AZR-Sonderauszählungsindividualdatensatzes,
- Sonderauszählungen zur Beschreibung der Grundgesamtheit,
- Stichprobenziehung in zwei Stufen (Ziehung der Ausländerbehörden, Ziehung der Personen),
- Bitte um Mitarbeit der Ausländerbehörden zur Zuspiegelung der Adresse zur Stichprobe auf Ausländerbehördenebene,
- Versand von Listen der Stichprobe an die Ausländerbehörden durch das BAMF zur Ergänzung der Listen um die Adresse,
- Einfügen der Adresse in die Liste und Rückversand durch die Ausländerbehörden, Vereinheitlichung und Bereinigung der Listen beim BAMF,
- Weitergabe der bereinigten Listen an das die Interviews durchführende Sozialforschungsinstitut,
- Bearbeitung der Liste und Verwendung der Liste durch das durchführende Sozialforschungsinstitut,
- Beginn der Durchführung der Interviews.

Allerdings sind auch bei der Ziehung einer bundesweiten Einwohnermeldeamtstichprobe in der Regel etwa fünf bis sechs Monate nötig, was die Dauer der Stichprobenziehung bei RAM 2006/2007 relativiert.²³ Auf eine grundsätzliche Schwierigkeit, die aber auch in ähnlicher Weise auf Einwohnermelderegister zutrifft, ist zudem hinzuweisen: Das AZR ist eine Prozessdatei. Einerseits wird es täglich mit neuen Daten, etwa von neu nach Deutschland zugewanderten Ausländern, beliefert. Andererseits sind andere Daten, wie etwa von einst ausländischen Personen, die die deutsche Staatsangehörigkeit erhielten, zu löschen. Je nach Schnelligkeit der Bearbeitung in den Ausländerbehörden, dem Bundesverwaltungsamt oder den weiteren zuliefernden und nutzenden Behörden kommt es unvermeidlich immer zu Verzögerungen und damit zu einer gewissen Fehlerfassung im AZR. Für RAM 2007/2007 bedeutet dies konkret: Ob der Abstand zwischen Stichprobenziehung und Befragung, in der zum Beispiel eine Zielperson die deutsche Staatsangehörigkeit erhielt, oder aber eine verzögerte Meldung und Löschung einer schon vor dem Stichtag deutschen Person im AZR dazu beitrug, dass letztlich rund 3% der Befragten zum Zeitpunkt der Befragung (auch) eine deutsche Staatsangehörigkeit hatten, kann letztlich nicht geklärt werden.

23 So wurde beim ALLBUS 2006 am 5. Oktober 2005 mit den Anschreiben an die Einwohnermeldeämter begonnen. Mit allen Nachfassaktionen dauerte es bis zum 23. Februar 2006, bis alle angeforderten Adressen aus den Gemeinden vorlagen, aus denen dann die Einsatzstichprobe gezogen wurde. Das Autorenteam dankt herzlich Martina Wasmer (GESIS, Mannheim) für diese Information.

Der Prozentsatz letztlich aus dem AZR verwertbarer Interviews ist bei RAM 2006/2007 ähnlich hoch wie jener mittels Einwohnermeldeämtern gewonnener Interviews beim GGS mit türkischen Zielpersonen. Die Abschätzung des Prozentsatzes realisierbarer Interviews kommt in beiden Fällen auf rund 28%. Vergleicht man dies mit dem ebenfalls über eine Einwohnermeldeamtsstichprobe realisierten ALLBUS 2006, in dem die deutliche Mehrheit der Zielpersonen jedoch Deutsche sind, dann stellt sich beim ALLBUS 2006 eine etwas bessere Verwertbarkeit der laut Einwohnermeldebehörden potenziellen Zielpersonen im Hinblick auf ein letztlich realisiertes Interview ein.

Auch wenn bei dem beschriebenen AZR-Verfahren nicht triviale Hürden zu überwinden sind, so gelang es schließlich doch, bei RAM 2006/2007, von 1.544 Türken, 972 Personen aus dem ehemaligen Jugoslawien, 746 Italienern, 677 Griechen und 637 Polen Daten zu vielfältigen integrations- und migrationsrelevanten Aspekten zu erheben. Damit liegt ein Datensatz vor, mit dem vertiefende Aussagen über die fünf größten Ausländergruppen in Deutschland möglich sind. Auch wenn man einen, mit Einwohnermeldeamtsstichproben etwa vergleichbaren, langen Vorlauf für eine Stichprobenziehung aus dem AZR benötigt, so hat sich doch gezeigt, dass man bei RAM 2006/2007 gezielter als bei den Vorgängeruntersuchungen an die Zielgruppe der Ausländer in Deutschland herantreten konnte. Zudem wurde die Erhebung mit der heute gängigen Vorgehensweise der Computer-unterstützten Befragung realisiert. Auch von diesem Gesichtspunkt aus ist mit einer gegenüber den Vorgängeruntersuchungen erhöhten Datenqualität auszugehen. Allerdings muss man sich mit dieser Vorgehensweise einer wichtigen Restriktion bewusst sein: Mit einer Stichprobenziehung aus dem AZR beziehen sich die Erkenntnisse, die mit einer Studie wie RAM 2006/2007 gewonnen werden, nur auf Ausländer im staatsrechtlichen Sinne, nicht aber auf weitere Personen mit Migrationshintergrund (Eingebürgerte, Personen mit doppelter Staatsangehörigkeit (deutsch-ausländisch) oder deutsche Nachfahren von Zuwanderern).

Da sich der Weg über eine Stichprobenziehung aus dem AZR aber letztlich als gangbar erwiesen hat, wird bei zukünftigen Untersuchungen des BAMF zu Ausländern zu prüfen sein, ob sich diese Vorgehensweise wieder zur Befragung anbietet.²⁴ Hierbei wird im Einzelfall abzuwägen sein, welche alternativen Verfahren unter welchen Kosten- und Zeitgesichtspunkten darüber hinaus zur Verfügung stehen. Sollen neben Ausländern auch Eingebürgerte, Personen mit doppelter Staats-

24 So wurde bei einer im Jahr 2007/2008 durchgeführten Befragung von hochqualifizierten Ausländern (Bundesamt für Migration und Flüchtlinge 2008b) ebenfalls das AZR als Datengrundlage zur Ermittlung der Befragten verwendet.

angehörigkeit oder auch deutsche Nachfahren von Zuwanderern befragt werden, sind zusätzliche Verfahren zur Gewinnung dieser Zielgruppen hinzuzuziehen.

Literatur

- Babka von Gostomski, C., 2007: Konzeption, Methoden und Durchführung der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“ (RAM). Stand: 30.10.2007. Unveröffentlicht. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Babka von Gostomski, C., 2008: Türkische, griechische, italienische und polnische Personen sowie Personen aus den Nachfolgestaaten des ehemaligen Jugoslawien in Deutschland. Erste Ergebnisse der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Brenke, K., 2008: Migranten in Berlin – Schlechte Jobchancen, geringe Einkommen, hohe Transferabhängigkeit. Wochenbericht DIW, Ausgabe 35/2008: (Berlin: DIW).
<http://www.diw-berlin.de/documents/publikationen/73/88442/08-35-1.pdf> (28.08.2008).
- Bundesamt für Migration und Flüchtlinge, 2007: Minas - Atlas über Migration, Integration und Asyl. (Nürnberg: Bundesamt für Migration und Flüchtlinge).
http://www.bamf.de/cln_092/nn_441592/SharedDocs/Anlagen/DE/Integration/Publicationen/Sonstige/atlas-minas-publikation,templatel=raw,property=publicationFile.pdf/atlas-minas-publikation.pdf (28.08.2008).
- Bundesamt für Migration und Flüchtlinge, 2008a: Ausländerzentralregister. (Nürnberg: Bundesamt für Migration und Flüchtlinge).
http://www.bamf.de/cln_011/nn_442016/SharedDocs/Anlagen/DE/Integration/Publicationen/Sonstige/atlas-minas-publikation,templatel=raw,property=publicationFile.pdf/atlas-minas-publikation.pdf (28.08.2008).
- Bundesamt für Migration und Flüchtlinge, 2008b: Qualifizierte Arbeitsmigration in Deutschland in wirtschafts-, arbeitsmarkt- und entwicklungspolitischer Hinsicht. (Nürnberg: Bundesamt für Migration und Flüchtlinge).
http://www.bamf.de/cln_092/nn_1026394/SharedDocs/Projekte/DE/Migration/Forschung/Migration/Arbeitsmigration/laufende/forschung-qualifizierte-arbeitsmigration-in-deutschland.html (28.08.2008).
- Bundesministerium des Innern und Bundesamt für Migration und Flüchtlinge, 2005: Migrationsbericht des Bundesamts für Migration und Flüchtlinge im Auftrag der Bundesregierung. Migrationsbericht 2005. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Bundesministerium für Arbeit und Soziales, 2008: Lebenslagen in Deutschland. Der 3. Armuts- und Reichtumsbericht der Bundesregierung. (Berlin: Bundesministerium für Arbeit und Soziales).
http://www.bmas.de/coremedia/generator/26742/property=pdf/dritter_armuts_und_reichtumsbericht.pdf (28.08.2008).
- Bundestag Drucksache 16/10155, 2008: Antwort der Bundesregierung auf die Große Anfrage der Abgeordneten Sibylle Laurischk, Ina Lenke, Miriam Gruß, weiterer Abgeordneter und der Fraktion der FDP – Drucksache 16/8301 – Seniorinnen und Senioren in Deutschland. Gesamtherstellung. Berlin: H. Heenemann GmbH & Co.
<http://dip21.bundestag.de/dip21/btd/16/101/1610155.pdf> (01.09.2008).
- Deutsche Forschungsgemeinschaft, 1999: Qualitätskriterien der Umfrageforschung. M. Kaase (Hg.): Denkschrift. Berlin: Akademie Verlag.
- Die Bundesregierung, 2007: Der Nationale Integrationsplan. Neue Wege – Neue Chancen. (Berlin: Presse- und Informationsamt der Bundesregierung).
<http://www.bundesregierung.de/Content/DE/Artikel/2007/07/Anlage/2007-10-18-nationaler-integrationsplan,property=publicationFile.pdf> (28.08.2008).

- Diehl, C., 2005: Der Integrationssurvey des Bundesinstituts für Bevölkerungsforschung. S. 11-22 in: S. Haug und C. Diehl (Hg.): Aspekte der Integration. Eingliederungsmuster und Lebenssituation italienisch- und türkischstämmiger junger Erwachsener in Deutschland. Wiesbaden: VS-Verlag für Sozialwissenschaften.
- Ette, A., G. Hullen, I. Leven und K. Ruckdeschel, 2007: Generations and Gender Survey. Dokumentation der Befragung von türkischen Migranten in Deutschland. (Wiesbaden: Bundesinstitut für Bevölkerungsforschung beim Statistischen Bundesamt). http://www.bib-demographie.de/cIn_051/nn_750454/SharedDocs/Publikationen/DE/Download/Materialienbaende/121b,templated=raw,property=publicationFile.pdf/121b.pdf (28.08.2008).
- Gabler, S. und S. Häder (Hg.), 2002: Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann.
- Granato, N., 1999: Die Befragung von Arbeitsmigranten: Einwohnermeldeamt-Stichprobe und telefonische Erhebung? ZUMA-Nachrichten 45: 44-60. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_45.pdf (27.11.2008).
- GESIS, 2008: Informationen zum Mikrozensus 2005 (Mannheim: GESIS). <http://www.gesis.org/dienstleistungen/daten/amtliche-mikrodaten/mikrozensus/grundfile/mz2005/allg-informationen/> (14.11.2008).
- Häder, S., 2000: Telefonstichproben. ZUMA How-to-Reihe, Nr. 6. (Mannheim: GESIS). http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/howto/how-to6sh.pdf (26.09.2008).
- Haug, S., 2008: Sprachliche Integration von Migranten in Deutschland. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Haug, S. unter Mitarbeit von S. Worbs und S. Theuer, 2005: Empirische Migrations- und Integrationsforschung im Referat 220. „Machbarkeitsstudie“. Unveröffentlicht. Arbeitspapier. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Humbert, A. und K. Schneiderheinze, 2000: Stichprobenziehung für telefonische Zuwandererumfragen. Einsatzmöglichkeiten der Namenforschung (Onomastik). ZUMA-Nachrichten 47: 36-64. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_47.pdf
- Humbert, A. und K. Schneiderheinze, 2002: Stichprobenziehung für telefonische Zuwandererumfragen. Praktische Erfahrungen und Erweiterung der Auswahlgrundlage. S. 187-208 in: S. Gabler und S. Häder (Hg.): Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann.
- Koch, A., 1998: Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. ZUMA-Nachrichten 42: 66-93. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_42.pdf
- König, P., G. Schultze und R. Wessel, 1986: Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der Bundesrepublik Deutschland. Repräsentativuntersuchung '85, Forschungsbericht 133 der Friedrich-Ebert-Stiftung. Bonn: Bundesministerium für Arbeit und Sozialordnung.
- Mehrländer, U., C. Ascheberg und J. Ueltzhöffer, 1996: Repräsentativuntersuchung '95: Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der Bundesrepublik Deutschland. Forschungsbericht 263 der Friedrich-Ebert-Stiftung. Berlin: Bundesministerium für Arbeit und Sozialordnung.
- Mehrländer, U., R. Hofmann, P. König und H.-J. Krause, 1981: Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der Bundesrepublik Deutschland. Repräsentativuntersuchung '80. Forschungsbericht 50 der Friedrich-Ebert-Stiftung. Bonn: Bundesministerium für Arbeit und Sozialordnung.

- Ministerium für Generationen, Familie, Frauen und Integration des Landes Nordrhein-Westfalen, 2008: Nordrhein-Westfalen – Land der neuen Integrationschancen. 1. Integrationsbericht der Landesregierung. (Düsseldorf: Ministerium für Generationen, Familie, Frauen und Integration des Landes Nordrhein-Westfalen).
http://www.mgffi.nrw.de/integration/PDFs/1_Integrationsbericht_25_09_2008.pdf (14.11.2008).
- Mohler, P. Ph., A. Koch und S. Gabler, 2003: Alles Zufall oder? Ein Diskussionsbeitrag zur Qualität von face to face-Umfragen in Deutschland. ZUMA-Nachrichten 53: 10-15.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_53.pdf (27.11.2008).
- Pupeter, M., 2007: Repräsentativbefragung: Ausgewählte Migrantengruppen in Deutschland 2006/2007 (RAM). Unveröffentlichter Methodenbericht. August 2007. München: TNS Infratest Sozialforschung.
- Salentin, K., 1999: Die Stichprobenziehung bei Zuwandererbefragungen. ZUMA-Nachrichten 45: 115-135.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_45.pdf (27.11.2008).
- Salentin, K., 2002: Zuwandererstichproben aus dem Telefonbuch: Möglichkeiten und Grenzen. S. 164-186 in: S. Gabler und S. Häder (Hg.): Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann.
- Salentin, K., 2007: Die Aussiedler-Stichprobenziehung. Methoden-Daten-Analysen 1 (1): 25-44.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/mda/Vol.1_Heft_1/MDA1_Salentin.pdf (27.11.2008).
- Salentin, K. und F. Wilkening, 2003: Ausländer, Eingebürgerte und das Problem einer realistischen Zuwanderer-Integrationsbilanz. Kölner Zeitschrift für Soziologie und Sozialpsychologie 55 (2): 278-298.
- Santacreu F. O., N. Rother und M. Braun, 2006: Stichprobenziehung für Migrantenpopulationen in fünf Ländern. Eine Darstellung des methodischen Vorgehens im PIONEUR-Projekt. ZUMA-Nachrichten 59: 72-88.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_59.pdf (27.11.2008).
- Schneekloth, U. und I. Leven, 2003: Woran bemisst sich eine „gute“ allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu den Hintergründen von Non-response in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS. ZUMA-Nachrichten 53: 16-57.
http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_53.pdf (27.11.2008).
- Schnell, R., P. Hill und E. Esser, 2005: Methoden der empirischen Sozialforschung. 7. Auflage. München: Oldenbourg.
- Schumann, S., 2000: Repräsentative Umfrage. Praxisorientierte Einführung in empirische Methoden und statistische Analyseverfahren. München: Oldenbourg.
- Seifert, W., 2007: Integration und Arbeit. Aus Politik und Zeitgeschichte 22-23/2007: 12-19.
- Seifert, W., 2008: Aussiedlerinnen und Aussiedler – neue Erfassungsmöglichkeiten und sozioökonomisches Profil. S. 11-23 in: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (Hg.): Statistische Analysen und Studien. Band 53. Düsseldorf: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen.
- Statistisches Bundesamt, 2007a: Bevölkerung und Erwerbstätigkeit. Ausländische Bevölkerung. Ergebnisse des Ausländerzentralregisters 2006. Fachserie 1 Reihe 2. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt, 2007b: Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus 2005. Fachserie 1 Reihe 2.2 Wiesbaden: Statistisches Bundesamt.

- Statistisches Bundesamt, 2008: Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus 2006. Fachserie 1 Reihe 2.2. Wiesbaden: Statistisches Bundesamt.
- Venema, M. und C. Grimm, 2002a: Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der Bundesrepublik Deutschland. Repräsentativuntersuchung 2001. Teil A: Türkische, ehemalige jugoslawische, italienische sowie griechische Arbeitnehmer und ihre Familienangehörigen in den alten Bundesländern und im ehemaligen West-Berlin. Berichtsband. Forschungsbericht im Auftrag des Bundesministeriums für Arbeit und Sozialordnung. (Offenbach: Bundesministerium für Arbeit und Sozialordnung). http://www.bamf.de/cln_042/nn_566316/SharedDocs/Anlagen/DE/Migration/Downloads/Sonstige/ram-studie-2001-berichtsband-teil-a,templateId=raw,property=publicationFile.pdf/ram-studie-2001-berichtsband-teil-a.pdf (28.08.2008).
- Venema, M. und C. Grimm, 2002b: Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der Bundesrepublik Deutschland. Repräsentativuntersuchung 2001. Teil A: Türkische, ehemalige jugoslawische, italienische sowie griechische Arbeitnehmer und ihre Familienangehörigen in den alten Bundesländern und im ehemaligen West-Berlin. Tabellenband. Forschungsbericht im Auftrag des Bundesministeriums für Arbeit und Sozialordnung. (Offenbach: Bundesministerium für Arbeit und Sozialordnung). http://www.bamf.de/cln_042/nn_566316/SharedDocs/Anlagen/DE/Migration/Downloads/Sonstige/ram-studie-2001-tabellenband-teil-a,templateId=raw,property=publicationFile.pdf/ram-studie-2001-tabellenband-teil-a.pdf (28.08.2008).
- Wasmer, M., E. Scholz und M. Blohm, 2007: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS 2006). ZUMA Methodenbericht 2007/09. (Mannheim: GESIS). http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2007/07_09_Wasmer.pdf (14.11.2008).
- Worbs, S., 2008: Die Einbürgerung von Ausländern in Deutschland. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Zühlke, S., 2008: Auswirkungen von Proxy-Interviews auf die Datenqualität des Mikrozensus. S. 3-10 in: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (Hg.): Statistische Analysen und Studien. Band 53. Düsseldorf: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen.

Korrespondenzadressen: Christian Babka von Gostomski
Bundesamt für Migration und Flüchtlinge
Referat 222 – Migrations- und Integrations-
forschung: Schwerpunkt Empirie
Frankenstr. 210
90461 Nürnberg
christian.babkavongostomski@bamf.bund.de

Monika Pupeter
TNS Infratest Sozialforschung GmbH
Landsberger Straße 338
80687 München
monika.pupeter@tns-infratest.com

On the Treatment of Non-Original Sample Members in the German Household Panel Study (SOEP)

Zur Behandlung von Stichproben-Neuzugängen im Sozio-ökonomischen Panel (SOEP)

*Martin Kroh, Rainer Pischner,
Martin Spieß and Gert G. Wagner*

Abstract

Household panel studies usually provide data on parents, children, and siblings, but not, e.g., on all ex-partners after a divorce or split in a cohabiting unit. This kind of information, however, can be an interesting source for intergenerational research or, more generally, for research on the dynamics of life course. To the best of our knowledge, among the diverse household panel studies, only the German Socio-Economic Panel (SOEP) contains data on individuals entering a sample household in or after wave two also after leaving that household again. This allows all household members to be traced once having lived in the sample household. In this paper, we discuss the rationale for tracing non-original sample members (Non-OSMs) in household panel studies. We also present results on the incidence and thus the relevance of Non-OSMs in the SOEP in general, and for specific research questions in particular.

Zusammenfassung

Haushalts-Panel-Datensätze stellen üblicherweise Daten für Eltern, Kinder und Geschwister zur Verfügung, nicht aber Daten aller Ex-Partner nach einer Scheidung oder nach der Trennung von Lebenspartnern. Diese Information stellt jedoch eine interessante Datenquelle für intergenerationale Untersuchungen beziehungsweise allgemeiner, für Untersuchungen zur Dynamik des Lebenslaufes dar. Unseres Wissens nach sind Daten über Personen, die in Welle zwei oder später in einen Panelhaushalt einziehen und diesen danach wieder verlassen, nur im Sozio-oekonomischen Panel (SOEP) verfügbar. Im SOEP werden alle Personen weiterverfolgt, die einmal in einem Panelhaushalt gelebt haben. Im vorliegenden Papier diskutieren wir die Idee, die den Weiterverfolgungsregeln bezüglich der Neuzugänge nicht-originale Stichprobenmitglieder des SOEP zugrunde liegt. Präsentiert werden weiterhin Häufigkeiten und damit die Relevanz dieser Stichprobenneuzugänge im SOEP im Allgemeinen und für einige Forschungsfragen im Speziellen.

1 Introduction

Prospective household panel studies like PSID in the US, BHPS in the UK, and the German SOEP were originally designed to give a picture of the composition of private households and the well-being of the members of those households at the point in time of the first survey and in the nearer future (e.g., Wagner et al. 1993, 2007). However, the endurance of some panel surveys, such as SOEP with its now 25 years, entails that a large percentage of the first-wave survey participants have now formed new households of their own with other individuals, many of whom were not sampled in the wave one households but joined survey participants' households later, either by birth or by moving in. Furthermore, many of the new individuals who grew into the survey after wave 1 have left these households since. These changes in the composition of households reflect typical dynamics of living arrangements over more than two decades, and it has become more and more apparent that the individual-level data from household panel surveys can be an excellent source for the analysis of these changing relationships.

From a practitioner's point of view, however, these changes prompt the not-so-trivial question of who should be followed up as a sample member of a longitudinal household panel study, and who should not. The main differences between panel studies exist in the tracing rules for adults who join the panel after wave 1. For example, household panel studies like the PSID do not follow individuals who entered after wave 1 and who leave the sampled household again. As an exception to this 'classic' rule, the BHPS follows new sample members if they are the parents of individuals sampled in wave 1. However, if parenthood is deemed to constitute an 'important relation' to an original sample member (short, OSM) in the BHPS (Jenkins/Siedler 2007), thus justifying the tracing of non-original sample members (short, non-OSM), then one must ask why a former marriage or partnership of an OSM to a non-OSM is not sufficiently 'important'.

Whereas all household panel studies provide data on parents, children, and siblings, there exists only one study to date that provides data on both ex-partners after a divorce or split in a cohabiting unit: namely, the SOEP (Jenkins/Siedler 2007: 6). This is due to the tracing rule adopted for the SOEP, whereby *all* household members aged 17 and older are traced after living in a sample household. This means that even individuals that entered the sample after the corresponding households joined the survey, i.e., non-OSMs, are followed up after they leave the originally sampled household.

Whereas the design of the random first-wave sample and also of later random refreshment samples is usually straightforward just as in any cross-sectional

survey, non-OSMs who join existing households by way of self-selection require adaptation of sampling concepts. These adaptations are reflected in the weighting concepts of a household panel study. In this paper, we report in more detail the tracing rules of SOEP and present results on the incidence and thus the relevance of non-original sample members after 25 years. Section 2 discusses the rationale for tracing all individuals from each sample household instead of original sample members only in household panel studies, and in particular in the SOEP. In Section 3, we briefly discuss the ramifications of the tracing of non-original sample members for the sampling design and thus the weighting scheme of the SOEP. Section 4 concludes.

2 Inclusion and Tracing Rules

In household panel surveys, the primary selected units are households, and the individuals living within these households are additional, implicitly selected units. Because households are 'artificial units', the question of which household and individual, respectively, should be a sample member in longitudinal household panels is not as trivial as it may at first seem. Whereas this question is easy to answer in the case of cross-sections and even cohort studies of individuals, the answer becomes much more difficult in the case of household panels. The difficulties arise from the fact that households are defined by a set of individuals, and this set may change over time. For example, let household h be a sample household at time t , consisting of one individual, i , who is a sample member. Suppose that an additional individual, j , moves into this household between t and $t+1$ together with a child. Further assume that individual j leaves household h together with the child between $t+3$ and $t+4$ to form a new household, k , still belonging to the same target population. In our example, individual i would be an original sample member (OSM), i.e., a unit selected and observed in the first wave and, correspondingly, individual j a non-OSM, i.e., a unit not observed in the first wave.

Thus, the set of OSM individuals includes all respondents of wave one, i.e. original participants, and their children living in the observed wave-one-household. Similarly, on an individual level, all household panel surveys consider children born into an OSM household after wave one, as well as children of OSM households living abroad at wave one to be original sample member. Non-OSM units at the household level are all existing not observed households in the population of interest as well as all existing households not in the population of interest. The population of interest, or 'target population', in the case of SOEP are the house-

holds in Germany and the population not of interest are households abroad. Even more general, one may add all not yet existing households to this set. At the individual level this set includes all persons not observed in OSM households, i.e., not-observed individuals existing in the population, those who are existing outside the target population, and even those that were not yet born in wave one outside the sample.

Let us assume that in the above example, t is the first wave. Then, household h is an OSM, but household k is not. If we define only the OSMs as the units to be traced, then only household h is a sample member at times $t = 1, 2, 3, 4$. In addition to the household itself, usually its members are of interest. Turning to the individual level, individual i is an OSM. In wave two, the non-original sample member j enters the household and, depending on the inclusion rule, he or she may or may not be a sample member in wave two and later waves. Usually, individual j would be interviewed in wave two and thus would be a sample member in this wave. However, the difference between the tracing rule adopted in the SOEP and rules adopted for other household panels is that in the SOEP, household k and individual j are sample members even in wave four when they no longer live with the OSM. In the same spirit, the child living in household k becomes a respondent at the age of 17 in the SOEP. The alternative, adopted by other surveys such as PSID and BHPS, would be to ignore household k and individual j (together with the child, who is not a biological child of person i) from wave three on.

The SOEP was started with the 'classic' tracing rule established by PSID that only those households – and all the eligible individuals living therein – are to be surveyed that are inhabited by OSM individuals as well as their children. The SOEP fieldwork followed this classic concept for the first six waves (1984–89). However, it turned out to be difficult for interviewers in the field to distinguish between OSMs and non-OSMs, and a great deal of interview data were deleted after collection. Beginning with wave seven in 1990, the decision was made by the SOEP group in Berlin and the fieldwork organization Infratest Sozialforschung¹ in Munich to survey all eligible individuals living in a sample household, and to follow those living in the sample household at time t from t to $t + 1$. Not eligible are all household members who explicitly refuse to participate and those who did not participate two times in succession. Note that these inclusion and tracing rules imply that even individuals may become sample members that never lived in an OSM household nor ever actually lived with an individual OSM.

1 To the best of our recollection, this was proposed by the head of the fieldwork organization, Bernhard von Rosenblatt.

Table 1 reports the frequency and the relative share of households in 2006 by their composition of individuals who were either members of the originally sampled households or were not part of the initially sampled SOEP households. This leads to three types of households: those populated exclusively by individuals who belong to the original sample of households, including children of those HH (here: OSM-HH), those populated both by individuals from original sample households *and* by individuals from non-sampled households (here: Mixed HH), and lastly, households populated only by individuals from originally non-sampled households (here: Non-OSM HH). Note that Table 1, according to our definition of OSM's in the panel context, treats not yet born children of OSMs and children of OSMs living abroad in wave one as OSMs themselves once they appear in their parents' households. Note also that samples A and B started in 1984, sample C in 1990, sample D in 1994/1995, sample E in 1998, sample F in 2000, sample G in 2002, and sample H in 2006.

Table 1 The Number of Households in SOEP 2006 by Sample and Non/OSM Status

Samples	Number of Households				Proportion of OSM Status in Percent		
	Total	OSM HH ^{*)}	Mixed HH	Non-OSM HH ^{**)}	OSM HH ^{*)}	Mixed HH	Non-OSM HH ^{**)}
A	2,821	1,572	950	299	0.557	0.337	0.106
B	655	392	223	40	0.599	0.340	0.061
C	1,717	1,123	461	133	0.654	0.268	0.078
D	222	150	68	4	0.676	0.306	0.018
E	686	567	96	23	0.826	0.140	0.034
F	3,895	3,394	450	51	0.871	0.116	0.013
G	859	786	69	4	0.915	0.080	0.005
H	1,506	1,506	-	-	1.000	1.000	-
All	12,361	9,490	2,317	554	0.768	0.187	0.045

Note. *) OSM-HHs are households with original sample members only. **) Non-OSM-HHs are households with non-original sample members only. Source. SOEP (Waves A to W).

Among the 12,361 households surveyed in 2006, more than 20% contain at least a single person not covered by the originally sampled households. This share steadily increases as a function of the age of samples: the most recent sample H from 2006, by definition, includes only originally sampled persons in each interviewed household. In samples F and G, drawn in 2000 and 2002, the OSM-HH reach a share of only 90% after seven and five waves, respectively. In the oldest samples, A and B, slightly more than half of the households are populated exclusively by respondents who were wave-one household members, children living abroad, and unborn

children of OSM's. Interestingly, more than 10% of the households in sample A of 2006 contain *no* individual who was part of the original sample of households. In the Appendix, we report the development of these different types of households in samples A through G separately.

Obviously, the number and proportion of household inhabited by non-OSM individuals varies over the samples, suggesting that the proportion of households with non-OSM individuals is increasing with the number of panel waves. Exceptions to this pattern are sample B, which is a sample where the household head did not have the German citizenship in 1984, and sample D, a sample of immigrants. Although the results of Table 1 suggest an increasing importance of non-OSMs, one may wonder whether these cases, from an individual perspective, represent a sustained enlargement of the SOEP survey or whether their proven mobility is indicative of a higher attrition rate in the following years. In other words, one may ask whether non-original sample members joining a SOEP household are more volatile than original sample members and thus, whether their contribution is only temporary. As the question of long-term participatory behaviour is difficult to address at the household level (since households can, in principle, switch status repeatedly between the OSM, Mixed, and Non-OSM types), we investigate the participatory behaviour of individual respondents (Table 2).

Table 2 distinguishes three groups of respondents. The first group is made up of those who were members of the originally sampled households *and* participated in the initial wave of each sample A through G in 1984, 1990, 1994/5, 1998, 2000, and 2002, $t_{0,A-G}$. The second group contains individuals who were members of the originally sampled households in samples A through G but were *not* interviewed in $t_{0,A-G}$. This applies to individuals who were too young to participate in $t_{0,A-G}$, who were not yet born or who were living abroad in wave one but became part of the active sample in one of the following years $t_{>0,A-G}$. The final group contains respondents who were not members of sampled households in $t_{0,A-G}$ and thus participated, like the second group, for the first time in $t_{>0,A-G}$, i.e., a wave subsequent to initial sampling. While the first two groups represent the raw sample of individuals who belong to the original sample of households, the third group covers external entrants to the survey.

Table 2 The Probability of Continued Participation of Persons by Non/OSM-Status

Years After First Interview	Original Sample Members		Non-Original Sample Members
	Participants in t_0	Non-Participants in t_0	
1	0.881	0.918	0.912
2	0.814	0.853	0.844
3	0.767	0.793	0.788
4	0.721	0.743	0.744
5	0.685	0.688	0.704
(...)			
10	0.563	0.477	0.530
(...)			
15	0.461	0.371	0.412
(...)			
20	0.379	0.274	0.311
N	35,899	5,268	6,275
Mean Age in t_0	44.75	19.03	29.60

Note. Entries denote Kaplan-Meier survival estimates of individual respondents' participation in the SOEP after their first interview. If respondents move abroad or die, we consider this event as a form of right-censoring. Source: SOEP (Waves A to W).

The entries in Table 2 denote the probability of continued participation in the SOEP after each individual's first interview. Note that this time point coincides with the year of the first waves of samples A through G only for the first group. Note also that the reported Kaplan-Meier survival estimates treat an exit from the survey due to moving abroad and death technically as a form of right-censoring, which does not affect the estimate of the probability of continued participation. The figures suggest that until the sixth wave of each individual's initial interview, new sample members have an even somewhat *higher* response probability, with 70%, than interviewees who already lived in the originally sampled households at $t_{0,A-G}$ with 69%. Only in the very long run is the continued participation of first-wave respondents better than in the two other groups. The latter is arguably due to the much lower age of the respondents who enter the SOEP in waves subsequent to initial sampling (mean age in t_0 of 19 and 30 respectively) as opposed to the sample of participating respondents in the initial wave (mean age in $t_{0,A-G}$ of 45 years). There is, however, no indication that new sample members are dramatically more volatile than those in the original sample of households.

3 The Advantage for Tracing Non-OSMs

In principle, a tracing rule that follows non-original sample members creates a kind of snowball-effect that, for the number of waves going to infinity and a population that is growing at a lower rate than the sample, would lead to the inclusion of all households and individuals living in the target population ('universe'). However, this does not happen at all due to the attrition rates of households and individuals. Much more importantly, the constant growing of non-original sample members into ongoing longitudinal household panel surveys may, in principle, capture some of the dynamics in the underlying population that cannot be reproduced efficiently in panel designs focusing exclusively on first-wave respondents.

As an example, consider the age of individuals. In a longitudinal survey of a fixed set of first-wave respondents ('original participants'), the average age will grow parallel with the length of the panel. Post-stratifying weighting variables would be the only way to compensate the growing difference between the age distribution in the sample and the underlying population. However, in household panel surveys, children of original sample members eventually reach the age of interviewing and thus grow into the sample, and individuals not sampled originally join existing households as well. In principle, these expansions of the sample diminish the differences in the age structure between the sample and the target population.

Table 3 presents the mean age of the subsamples of the SOEP samples A to H: first, it gives the mean age of persons in 2007 who were already interviewed in the first wave of the respective samples. As one would expect, the more distant the year of sampling, the higher the mean age of respondents. Due to age differences between subsamples already existing in the years of sampling, to deaths, and to age-specific attrition rates, the mean age of respondents does not precisely mirror the 'age' of the subsamples. Second, the table shows the mean age of respondents in 2007 who were sampled in the first waves of each subsample but who did not provide an interview. In most cases, these are persons who were 16 or younger in the first wave and grew into the survey in later years. This also includes not yet born children of original sample members. Obviously, the mean age of this group in 2007 is much lower than in the first group. Taking the average of both groups of original sample members shows that the inclusion of these first-wave respondents who did not provide an interview in the first wave significantly reduces the mean age of older samples such as sample A and B of 1984 by about ten years.

Third, the table gives the mean age of respondents in 2007 who did not provide an interview in the first waves of each subsample and who were not sampled at all in wave-one. This group thus consists mostly of individuals who entered

the SOEP sample by moving into an existing household. The mean age of this group in 2007 is between that of the first two groups, and reduces the average age of the cross-section by another two to three years in the oldest samples A and B of 1984 and also C of 1990.

Table 3 The Mean Age of Observed SOEP Respondents in 2007 by Sample and Type of Entrance in the Survey

Sample	Wave	Gross Sample in t_0 and children			Mean Age of Non-OSM's	Total Mean Age
		Mean Age of OSM's with Interview in t_0	Mean Age of OSM's without Interview in t_0	Mean Age of All OSM's. Average of Columns 3 and 4		
A	24	61	28	52	41	49
B	24	57	29	45	38	43
C	18	57	26	49	37	47
D	14	50	22	46	37	45
E	12	55	22	51	38	50
F	8	54	25	51	36	50
G	6	53	24	50	35	49
H	2	51	33	51	34	51
Total		55	26	50	39	49

In the end, the mean age of respondents in all SOEP subsamples in 2007 hovers around 50 years, although samples A and B were drawn more than 20 years prior to the most recent refreshment sample H. The descriptive results in Table 3 thus imply that excluding the non-OSM individuals, i.e., those who enter the sample after wave one, leads to higher mean ages in the sample than in the 'augmented' samples, where these non-OSM individuals are included.

Obviously, the results in Table 3 suggest that ignoring the non-OSMs may lead to selective samples with respect to characteristics of the underlying cross-sectional population, at least with respect to age and variables correlated with age. Ignoring these units would thus require additional modelling of the corresponding missing mechanism at the analysis stage. These corrections usually are performed – as in BHPS and PSID – by way of post-stratified weighting. The inclusion of non-OSMs in principle reduces the variance of cross-sectional weights and thus increases the efficiency of any weighted estimation. Note that for some dynamics in the underlying population, such as new inflows of migrants, it is unlikely that the inclusion of non-original sample members sufficiently captures these developments. This due to the fact that the fusion of households is likely to be a non-

random process, but one that is selective with respect to household characteristics. Since it is rather unlikely that immigrants, who form new households and thus become relevant population units, are affected by this strategy in a rather small number of waves, the German SOEP accounts for immigrants that do not enter existing households through its special 'immigrant sample', sample D (cf. Burkhauser et al. 1997) and through general refreshment samples (cf. Wagner et al. 2007).

Besides the advantage that augmenting the sample with non-original sample members reduces the need for weighting, the broad tracing rules in SOEP facilitate a number of specific research questions that focus on relationships ('dyads'). One prominent example is the intergenerational transmission literature, e.g., the transmission of poverty from parents to children (Jenkins/Siedler 2007), the well-being of widows and widowers (Burkhauser et al. 2005; Lucas 2007), and the impact of children's well-being on the happiness of their parents (Schwarze/Winkelmann 2005).

Table 4 illustrates that a large portion of the data that permits the matching of individual respondents with their parents, siblings, but also grandparents, uncles, and aunts involve non-original sample members. For instance, more than half of the individuals with information on their parents were not interviewed in wave one, most because they were too young (16 years or younger). Restricting the units of analysis to first-wave respondents clearly reduces the number and the types of relationships.

Table 4 Shares of Respondents with Interviews on Relatives in the SOEP and Total Figures on Respondents by the Type Entrance in the Survey

	Original Sample Member		Non-Original Sample Members	Total	
	Person with Interview in t_0	Person without Interview in t_0		All	N
Parents	0.42	0.52	0.06	1.00	11,366
Siblings	0.34	0.60	0.06	1.00	8,394
Grandparents	0.25	0.59	0.16	1.00	729
Uncles/Aunts	0.02	0.75	0.23	1.00	238
Total	0.75	0.13	0.12	1.00	48,016

Note. Reading example: Of the 11,366 SOEP respondents interviewed having either a father or mother who was also interviewed, 42% were already interviewed in wave one of the respective SOEP sample, 52% were part of the wave one gross sample but did not participate in wave-one, and 6% entered the gross sample after wave one. *Data Source.* SOEP data distribution 2008.

Further exciting research questions have recently been opened up by the possibilities of linked data – not only on parents and children but also on the linked life courses, e.g., of siblings and couples (Ermish et al. 2006). Linked life courses allow analyses along the lines of the 'behavioural genetics' approach, which attempts to disentangle the impacts of nature ('genes') and economic respectively social circumstances ('environment') on human behaviour and well-being. Schimmack/Lucas (2007), for example, used SOEP data to analyze the well-being of the same couples during marriage and after divorce.

Table 5 reports the incidence of partnership dissolutions in the SOEP for which data exists on both partners before *and* after the separation. About half of these cases would not have been possible to observe if the tracing rules had covered first-wave respondents only. SOEP's tracing rules allow partnership dissolutions to be observed more than once in several individuals: without following up both original and non-original sample members, particularly the consequences of such repeated separations would be impossible to investigate.

Table 5 Dissolution of Partnerships with Interviews with both Partners before and after Separation by Type of Entrance into the Survey

	Both Persons	Mixed Couples	Both Persons	Total	
	Original Sample Members		Non-Original Sample Members	All	N
1 st Dissolution for Both Partners	0.47	0.45	0.06	1.00	838
2 nd Dissolution for at least one Partner	0.43	0.49	0.08	1.00	357
3 rd + Dissolution for at least one Partner	0.19	0.68	0.13	1.00	69

As yet another example, consider the living conditions, income, and well-being of widows/widowers. Every widow/widower in the sample, whether an OSM individual or not, can contribute to this kind of research. Over the course of time in a long-running panel, the percentage of non-OSM widows/widowers in the sample increases because more and more widows/widowers enter the sample through marriage to OSM individuals. Ignoring these non-OSM individuals would lead to a loss of a great amount of information, for example, about the 'final resting place' of the deceased OSM individuals (cf. Gerstorff et al. 2008), a topic that has received increasing attention in recent years.

4 Implications of Inclusion Rules for Weighting

The choice of inclusion and tracing rules not only is important for facilitating substantive research, but clearly also has ramifications for the weighting scheme. Weights as usually delivered with archived data sets, are generated to allow users of the data sets to compensate for different selection and response probabilities. At the heart of most theoretically justified weighting approaches is the probability of observing the corresponding unit at the corresponding point in time (e.g. Robins/Rotnitzky/Zhao 1995; Wooldridge 2002). Selection probabilities are usually known from the sampling design, whereas response probabilities have to be estimated. Obviously, units are only observed if they are selected into the sample and if they participate. The observation probabilities can be modelled by a sequence of known or estimated conditional probabilities. In principle, this is the procedure adopted for the SOEP (Kroh/Spieß 2008; Goebel et al. 2008: 88–96). It allows a straightforward incorporation of Non-OSM in the weighting scheme.

If the rule is adopted to trace all eligible individuals inhabiting a sample household from wave t to wave $t+1$ and to survey all households in $t+1$ inhabited by such a former-wave sample individual during that period of time, i.e., if we want to also include non-original sample members once they form a new household, the task is to derive, model, and estimate the probabilities of observing households that were observed in t and $t+1$ and, for new households, of observing households that were not observed in t but are observed in $t+1$. Modelling and estimating the observation probabilities of households with move-ins from t to $t+1$ can be demanding, depending on whether the new household member was (a) part of the sample, (b) part of the population of interest, i.e., living in Germany, but not in the sample, and (c) not part of the population in t , i.e., living abroad.

However, deriving the observation probabilities for split-off households is, under assumptions, generally straightforward. Without additional nonresponse on the individual level, the probability of observing individuals from a sample household is equal to the probability of observing that particular household (Galler 1987).

One requirement for estimating observation probabilities for a variety of groups is that the number of cases in the sample be sufficient to allow the estimation of group-specific effects on response behaviour. The smaller the number of observations within relevant groups with different response behaviours becomes over time, the more one has to rely on the models used to generate the weights. Furthermore, ignoring cases in $t+1$ that are in the sample at time t introduces a missing data problem. If these missing cases (households or individuals) are not missing completely at random (e.g., Rubin 1987), then the missing data process cannot gen-

erally be ignored. Thus, the inclusion respectively the tracing rules adopted in the SOEP to include and trace all eligible household members prevents modelling and estimation of this aspect of the (non-)response models, which in turn are part of the process of deriving or estimating observation probabilities. Again, the efficiency of the weighting scheme benefits from the broad inclusion rules.

This point is illustrated in Table 3, where inclusion of those individuals that entered the sample later than wave one lowered the mean age of most of the samples to the level of the latest sample, sample H, selected in 2006. At least part of a possible selectivity with respect to age and variables correlated with age can be overcome by allowing these units to grow into the sample. Thus, it is not necessary to model this aspect of the (non-)response process, relying on (untested) assumptions.

5 Conclusions

The German Socio-Economic Panel Study (SOEP) is a household panel survey with a different tracing rule than other household panel studies like PSID and BHPS. PSID drops Non-OSMs once the OSM leaves a sample household. BHPS does so as well, but with the exception of those Non-OSMs who are parents of OSM-children. Only SOEP traces all Non-OSMs. That is, even if a new, i.e., non-original household member leaves the sample household, the SOEP considers this individual an established part of the survey and continues to trace her subsequent living arrangements. This strategy, originally adopted on a non-theoretical basis to make fieldwork more efficient, turned out to enrich the data significantly because it reduces the necessity to weight and opens up new research questions.

If the classical tracing rule is applied, the analyst has not only to compensate (if necessary) for differing sampling probabilities, first-wave nonresponse and attrition in later waves to derive valid inferences, but also for ignoring non-OSMs. Given the results reported in Section 2, these missing units must be assumed to be not missing completely at random, which requires appropriate weighting. A simple strategy, however, is to trace and include these units into the sample.

Non-OSMs will eventually make up a large portion of the respondents and households in any long-running panel design. Analyses reported in this paper do not suggest that these cases are more volatile in their participatory behaviour than OSMs. Furthermore, non-OSMs allow researchers to address innovative research questions. It has been clear since the very beginning that following Non-OSMs can be helpful in analyzing the impact of events like divorces or separations of cohabiting units. However, it has just recently been shown that the increased number of

cases providing data on respondents who have lived together for some time and split up is extremely valuable for disentangling the influence of genes and environment based on the differences in biological and (changing) social factors (cf. Schimmack/Lucas 2007). In addition, tracing Non-OSMs is helpful for the analysis of the terminal phase of life (cf. Gerstorff et al. 2008). The value of this tracing rule will increase further after the introduction of 'exit interviews' into the SOEP dealing with the terminal phase of life of respondents who have passed away (Kröger 2008).

6 References

- Burkhauser, R. V., P. Giles, D. R. Lillard and J. Schwarze, 2005: Until Death Do Us Part: An Analysis of the Economic Well-Being of Widows in Four Countries. *Journals of Gerontology Series B- Psychological Sciences and Social Sciences* 60 (5): 238–246.
- Burkhauser, R. V., M. Kreyenfeld and G. G. Wagner, 1997: The German Socio-Economic Panel: A representative sample of reunited Germany and its parts. *Vierteljahrshefte zur Wirtschaftsforschung* 66 (1): 7–15.
- Ermish, J., M. Francesconi and Thomas Siedler, 2006: Intergenerational Mobility and Marital Sorting. *The Economic Journal* 116: 659–679.
- Galler, H. P., 1987: Zur Längsschnittgewichtung des Sozio-ökonomischen Panels. S. 295–317 in: H.-J. Krupp and U. Hanefeld (eds.): *Lebenslagen im Wandel*. Campus: Frankfurt and New York.
- Gerstorff, D., N. Ram, R. Estabrook, J. Schupp, G. G. Wagner and U. Lindenberger, 2008: Life Satisfaction Shows Terminal Decline in Old Age. Longitudinal Evidence from the German Socioeconomic Panel Study. *Developmental Psychology* (in press).
- Goebel, J., M. Grabka, P. Krause, M. Kroh, R. Pischner, I. Sieber and M. Spieß, 2008: Mikrodaten, Gewichtung und Datenstruktur der Längsschnittstudie Sozio-ökonomisches Panel (SOEP). *Vierteljahrshefte zur Wirtschaftsforschung* 77: 77–109.
- Haisken-DeNew, J. P. and J. R. Frick, 2005: DTC Desktop Companion to the German Socio-Economic Panel (SOEP). Version 8.0 - Updated to Wave 21. (Berlin: DIW). <http://www.diw.de/documents/dokumentenarchiv/17/38951/dtc.354256.pdf> (7.11.2008)
- Jenkins, S. P. and T. Siedler, 2007: Using household panel data to understand the intergenerational transmission of poverty. DIW Discussion Paper 694.
- Kröger, K., 2008. "Exit-Interviews" in SOEP und SHARE. DIW Data Documentation (forthcoming).
- Kroh, M. and M. Spieß, 2008: Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP 1984 until 2007). DIW Data Documentation 39.
- Lucas, R. A., 2007: Adaption and the Set-Point Model of Subjective Well-Being. *Current Directions in Psychological Science* 16 (2): 75–79.
- Rendtel, U., 1995: *Lebenslagen im Wandel: Panelfälle und Panelrepräsentativität*. Frankfurt/New York: Campus.
- Robins, J. M., A. Rotnitzky and L. P. Zhao, 1995: Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association* 90 (429): 106–121.
- Rubin, D. B., 1987: *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.
- Särndal, C.-E., B. Swensson and J. Wretman, 1992: *Model Assisted Survey Sampling*. New York: Springer.
- Schimmack, U. and R. Lucas, 2007: Marriage Matters: Spousal Similarity in Life Satisfaction. *Schmollers Jahrbuch* 127 (1): 105–111.

- Schwarze, J. and R. Winkelmann, 2005: What can happiness research tell us about altruism? Evidence from the German Socio-Economic Panel. DIW Discussion Paper 475.
- Wagner, G. G., R. V. Burkhauser and F. Behringer, 1993: The English Language Public Use File of the German Socio-Economic Panel. The Journal of Human Resources 28 (2): 429-433.
- Wagner, Gert G., J. R. Frick, and J. Schupp, 2007: The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. Schmollers Jahrbuch 127 (1), 139-169.
- Wooldridge, J. M., 2002: Econometric Analysis of Cross Section and Panel Data. Cambridge/Massachusetts: The MIT Press.

Appendix

Table A1-A Status of SOEP Households in Sample A
- Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
1984(A)	4,528	4,528	-	-	100.0	-	-
1985(B)	4,141	3,968	173	-	95.8	4.2	-
1986(C)	3,962	3,643	319	-	91.9	8.1	-
1987(D)	3,910	3,509	401	-	89.7	10.3	-
1988(E)	3,743	3,279	463	1	87.6	12.4	0.0
1989(F)	3,647	3,093	552	2	84.8	15.1	0.1
1990(G)	3,612	2,933	654	25	81.2	18.1	0.7
1991(H)	3,613	2,833	738	42	78.4	20.4	1.2
1992(I)	3,585	2,723	798	64	76.0	22.3	1.8
1993(J)	3,603	2,657	842	104	73.7	23.4	2.9
1994(K)	3,577	2,541	896	140	71.0	25.1	3.9
1995(L)	3,526	2,417	945	164	68.6	26.8	4.6
1996(M)	3,485	2,333	967	185	66.9	27.8	5.3
1997(N)	3,458	2,240	1,011	207	64.8	29.2	6.0
1998(O)	3,387	2,154	1,016	217	63.6	30.0	6.4
1999(P)	3,325	2,055	1,040	230	61.8	31.3	6.9
2000(Q)	3,240	1,984	1,016	240	61.2	31.4	7.4
2001(R)	3,168	1,896	1,015	257	59.9	32.0	8.1
2002(S)	3,123	1,847	1,010	266	59.1	32.3	8.5
2003(T)	3,072	1,784	1,009	279	58.1	32.8	9.1
2004(U)	3,010	1,718	1,004	288	57.1	33.4	9.6
2005(V)	2,937	1,655	977	305	56.3	33.3	10.4
2006(W)	2,821	1,572	950	299	55.7	33.7	10.6
Total (A to W)	80,473	59,362	17,796	3,315	73.8	22.1	4.1

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.

Sources: SOEP (Waves A to W); author's calculations.

Table A1-B Status of SOEP Households in Sample B
 - Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
1984(A)	1,393	1,393	-	-	100.0	-	-
1985(B)	1,181	1,170	11	-	99.1	0.9	-
1986(C)	1,128	1,085	43	-	96.2	3.8	-
1987(D)	1,116	1,030	86	-	92.3	7.7	-
1988(E)	1,071	952	119	-	88.9	11.1	-
1989(F)	1,043	889	152	2	85.2	14.6	0.2
1990(G)	1,028	852	167	9	82.9	16.2	0.9
1991(H)	1,056	848	189	19	80.3	17.9	1.8
1992(I)	1,060	828	212	20	78.1	20.0	1.9
1993(J)	1,064	814	227	23	76.5	21.3	2.2
1994(K)	1,023	763	233	27	74.6	22.8	2.6
1995(L)	982	717	238	27	73.0	24.2	2.8
1996(M)	960	676	252	32	70.4	26.3	3.3
1997(N)	931	637	259	35	68.4	27.8	3.8
1998(O)	898	607	249	42	67.6	27.7	4.7
1999(P)	858	570	246	42	66.4	28.7	4.9
2000(Q)	820	532	249	39	64.9	30.4	4.8
2001(R)	809	519	250	40	64.2	30.9	4.9
2002(S)	766	483	244	39	63.1	31.9	5.1
2003(T)	742	462	237	43	62.3	31.9	5.8
2004(U)	714	431	240	43	60.4	33.6	6.0
2005(V)	698	418	237	43	59.9	34.0	6.2
2006(W)	655	392	223	40	59.8	34.1	6.1
Total (A to W)	21,996	17,068	4,363	565	77.6	19.8	2.6

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.

Sources: SOEP (Waves A to W); author's calculations.

Table A1-C Status of SOEP Households in Sample C
- Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
1990(G)	2,179	2,179	-	-	100.0	-	-
1991(H)	2,030	1,968	61	1	97.0	3.0	0.0
1992(I)	2,020	1,888	125	7	93.5	6.2	0.3
1993(J)	1,970	1,792	167	11	90.1	8.5	0.6
1994(K)	1,959	1,701	237	21	86.8	12.1	1.1
1995(L)	1,938	1,635	272	31	84.4	14.0	1.6
1996(M)	1,951	1,602	311	38	82.1	15.9	2.0
1997(N)	1,942	1,549	339	54	79.8	17.	2.8
1998(O)	1,886	1,485	345	56	78.4	18.3	3.0
1999(P)	1,894	1,458	366	70	77.0	19.3	3.7
2000(Q)	1,879	1,409	397	73	75.0	21.1	3.9
2001(R)	1,850	1,367	399	84	73.9	21.6	4.5
2002(S)	1,818	1,321	404	93	72.7	22.2	5.1
2003(T)	1,807	1,256	437	114	69.5	24.2	6.3
2004(U)	1,813	1,209	479	125	66.7	26.4	6.7
2005(V)	1,771	1,182	460	129	66.7	26.0	7.3
2006(W)	1,717	1,123	461	133	65.4	26.8	7.8
Total (G to W)	32,424	26,124	5,260	1,040	80.6	16.2	3.2

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.

Sources: SOEP (Waves G to W); author's calculations.

Table A1-D Status of SOEP Households in Sample D
- Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
1995(L)	322	316	6	-	98.1	1.9	-
1996(M)	302	287	15	-	95.0	5.0	-
1997(N)	286	259	27	-	90.6	9.4	-
1998(O)	259	224	35	-	86.5	13.5	-
1999(P)	252	202	49	1	80.2	19.4	0.4
2000(Q)	249	197	48	4	79.1	19.3	1.6
2001(R)	234	182	51	1	77.8	21.8	0.4
2002(S)	244	177	64	3	72.5	26.2	1.2
2003(T)	248	176	67	5	71.0	27.0	2.0
2004(U)	236	165	66	5	69.9	28.0	2.1
2005(V)	233	155	75	3	66.5	32.2	1.3
2006(W)	222	150	68	4	67.6	30.6	1.8
Total (L to W)	3,087	2,490	571	26	80.7	18.5	0.8

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.

Sources: SOEP (Waves L to W); author's calculations.

Table A1-E Status of SOEP Households in Sample E
- Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
1998(O)	1,056	1,056	-	-	100.0	-	-
1999(P)	886	862	24	-	97.3	2.7	-
2000(Q)	838	793	43	2	94.6	5.1	0.2
2001(R)	811	745	60	6	91.9	7.4	0.7
2002(S)	773	689	73	11	89.1	9.4	1.4
2003(T)	744	646	83	15	86.8	11.2	2.0
2004(U)	732	623	93	16	85.1	12.7	2.2
2005(V)	706	593	95	18	84.0	13.5	2.6
2006(W)	686	567	96	23	82.7	14.0	3.4
Total (O to W)	7,232	6,574	567	91	90.9	7.8	1.3

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.

Sources: SOEP (Waves O to W); author's calculations.

Table A1-F Status of SOEP Households in Sample F
- Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
2000(Q)	6,052	6,052	-	-	100.0	-	-
2001(R)	4,911	4,796	115	-	97.7	2.3	-
2002(S)	4,586	4,380	200	6	95.5	4.4	0.1
2003(T)	4,386	4,081	295	10	93.1	6.7	0.2
2004(U)	4,235	3,836	373	26	90.6	8.8	0.6
2005(V)	4,070	3,613	415	42	88.8	10.2	1.0
2006(W)	3,895	3,394	450	51	87.1	11.6	1.3
Total (Q to W)	32,135	30,152	1,848	135	93.8	5.8	0.4

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.
Sources: SOEP (Waves Q to W); author's calculations.

Table A1-G Status of SOEP Households in Sample G
- Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
2002(S)	998	998	-	-	100.0	-	-
2003(T)	911	889	22	-	97.6	2.4	-
2004(U)	902	865	36	1	95.9	4.0	0.1
2005(V)	879	827	48	4	94.1	5.5	0.5
2006(W)	859	786	69	4	91.5	8.0	0.5
Total (S to W)	4,549	4,365	175	9	96.0	3.8	0.2

*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.
Sources: SOEP (Waves S to W); author's calculations.

Table A1-H Status of SOEP Households in Sample H
 - Households by Year (Wave) and OSM Status -

	Households Total	OSM-Households*)	Mixed Households	NOSM-Households**)	OSM-Households*)	Mixed Households	NOSM-Households**)
Year(Wave)	Number of Households				Shares of OSM Status in percent		
2006(W)	1,506	1,506	-	-	100.0	-	-

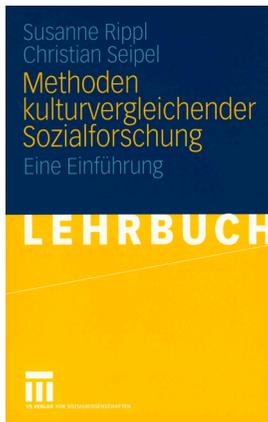
*) OSM-Households: households with original sample members only **) NOSM-Households: households with non-original sample members only.

Sources: SOEP (Wave W); author's calculations.

Contact Address:

Martin Kroh
 Deutsches Institut für Wirtschaftsforschung
 Abteilung SOEP
 Mohrenstraße 58
 D-10117 Berlin
 mkroh@diw.de

Rezensionen



SUSANNE RIPPL und
CHRISTIAN SEIPEL,
2008: Methoden
kulturvergleichender
Sozialforschung.
Eine Einführung.
Wiesbaden: VS
Verlag für Sozial-
wissenschaften.
ISBN 978-3-3531-
14965-3, 189 Seiten,
19,90 EUR.

Zwei der vier berühmten Gesetze der Soziologie von Esping-Andersen lauten: „Everything is different in the South“ und „Nothing ever works in India“. Wenn wir nicht wissen, weshalb das der Fall sein sollte, dann gibt es immer die ‚Kultur‘, die im Süden anders ist als im Norden und in Indien anders als in Kiribati. In der vergleichenden Sozialforschung wird dieses Prinzip nicht nur als faule Ausrede verwendet, um unser beschränktes Wissen hinter prächtigen Begriffen zu verstecken. Man kann es auch als wichtigste Herausforderung eben dieser Forschung betrachten, den Kulturbegriff empirisch zugänglich zu machen und seinen Auftritt als Deus ex Machina zu vermeiden. Auch Esping-Andersens Gesetze sollten falsifizierbar sein.

Methodische Auseinandersetzungen mit den Problemen der vergleichenden Forschung sind heutzutage meistens sehr avanciert und die rasant wachsenden Datenberge bieten ideale Spielplätze für die übliche Kraftmeierei der Techniker. Für Sozialwissenschaftler, die zum Beispiel Kinderwünsche, Wählerverhalten, religiöse Institutionen oder Kriminalitätsraten unter verschiedenen Gruppen vergleichen möchten, ist das alles selten hilfreich. In ihrem neuen Lehrbuch stellen Susanne Rippl und Christian Seipel sich die

Aufgabe, „... eine praxisorientierte Einführung in die Methoden kulturvergleichender Forschung vorzulegen“, wobei der Schwerpunkt auf ‚Surveyforschung‘ liegt. Für Studierende und wenig erfahrene Forscher ist das sicherlich eine sehr willkommene Handreichung.

Das Buch umfasst neben einem als ‚Einleitung‘ bezeichnetem Vorwort und einem sehr mageren Schlusskapitel von nur einer Seite insgesamt sieben Kapitel. Auf der Basis vieler Beispiele aus verschiedenen Disziplinen beschreiben die Autoren den Forschungsprozess schrittweise und erläutern insbesondere die Komplikationen bei der Erstellung von Äquivalenz. Zunächst stellen Rippl und Seipel sich in Kapitel 2 die Frage, „Why culture matters“ und diskutieren die verschiedenen Kulturdefinitionen. In der Umfrageforschung hat sich ein subjektbezogener Kulturbegriff durchgesetzt und werden Kulturvergleiche faktisch als ländervergleichende Studien durchgeführt. Rippl und Seipel legen sich am Anfang zu Recht nicht auf derartige Spezialfälle fest, sondern bieten systematische Überblicke der unterschiedlichen Begriffsverwendungen und Forschungspraxen. Diese systematische Vorgehensweise wird in Kapitel 3 mit einer Darstellung der ‚methodologischen Debatten‘ in verschiedenen Disziplinen fortgesetzt. Beispiele aus der Ethnologie, Soziologie, Politikwissenschaft und Psychologie machen die unterschiedlichen Fragestellungen und Ansätze deutlich. Um diese Differenzen zu erläutern, vermeiden die Autoren grundlegende Auseinandersetzungen mit epistemologischen Fragen der Sozialforschung nicht. Eine multi-disziplinäre und für Anfänger verständliche Darstellung dieser Thematik ist natürlich extrem schwierig. Rippl und Seipel meistern diese Aufgabe mit wechselndem Erfolg und brauchen am Ende des Kapitels ein sehr langes ‚Fazit‘, um die vielen Punkte einigermaßen überschaubar zusammenzufassen und sich zu einer

„kritisch-rationellen Position“ zu bekennen. Ihre disziplinären Vertiefungen in diesem Kapitel sind unausgewogen und insbesondere die politikwissenschaftliche Kulturforschung wird mit nur einigen Absätzen abgetan. Von der außergewöhnlich einflussreichen Studie von Almond und Verba erfahren die Leser zum Beispiel kaum mehr, als dass es sich dabei um einen „Klassiker“ handelt. Max Kaases kritische Darstellung der politischen Kulturforschung als der Versuch, einen Pudding an die Wand zu nageln, wird nicht einmal erwähnt.

Nachdem die Grundlagen der vergleichenden Kulturforschung in den ersten drei Kapiteln dargestellt sind, widmen Rippl und Seipel sich in den weiteren Kapiteln dem praktischen Ablauf des Forschungsprozess bei standardisierten Umfragen. Zuerst betonen sie in Kapitel 4, dass das „Postulat der Vergleichbarkeit“ in allen Phasen kulturvergleichender Studien – also von der Problemdefinition bis zur Publikation – zu berücksichtigen ist. Diese Behandlung des Gesamtprozesses ist sehr wichtig, da viele Forscher Vergleichbarkeit noch immer als ein Problem der Operationalisierung und Messung betrachten. In diesem Kapitel präsentieren Rippl und Seipel die Herstellung von Äquivalenz als die zentrale Herausforderung der vergleichenden Forschung: „Nur gleichwertiges kann auch verglichen werden“. Anhand der zwei grundlegenden Fragen „What to measure?“ und „How to measure?“ werden die verschiedenen Äquivalenzbegriffe vorgestellt und mit Beispielen veranschaulicht. Auf diese Weise entsteht ein klarer Überblick über die Äquivalenzproblematik sowie die Quellen möglicher Verzerrungen. Und weil es so viele unterschiedliche Fehlerquellen gibt, ist jeder Kulturvergleich wegen der möglichen Kumulation von Verzerrungen stark bedroht.

Auswahlverfahren und Fragebogenentwicklung bilden die Themen von Kapitel 5 bzw. 6. Obwohl Rippl und Seipel in früheren Kapiteln zu Recht vor der einfachen Gleichsetzung von Kulturen, Gesellschaften und

Ländern warnen, beschäftigt sich Kapitel 5 ausschließlich mit Problemen der Auswahl von Ländern und Individuen. Merkwürdiger Weise sind sie dabei so auf Stichprobenziehungen fokussiert, dass sie einfach übersehen, dass für manche Fragestellungen Vollhebungen vorhanden sind (zum Beispiel in Lijpharts Studie demokratischer Staaten oder in Putnams Vergleich italienischer Regionen). Kapitel 6 diskutiert die allgemeinen Probleme der Anwendung vorhandener Instrumente, die Komplikationen bei Übersetzungen sowie die Antwortverzerrungen bei Fragebogenentwicklungen. Bereits in den ersten Kapiteln sehen Rippl und Seipel sich gezwungen, ihre Darstellungen des kulturvergleichenden Forschungsprozesses mit ausführlichen Erläuterungen über eher allgemeine Aspekte der Umfrageforschung zu ergänzen. Für ein Lehrbuch ist es sicherlich kein Nachteil, dass solche Sachen hier nochmals von einer vergleichenden Perspektive beleuchtet werden. Allerdings rücken die Darstellungen dieser allgemeinen Aspekte die spezifischen Probleme der vergleichenden Kulturforschung in Kapitel 5 und 6 immer mehr in den Hintergrund. Der Leser erhält so zwar nebenbei einen Auffrischkurs „Grundlagen der Umfrageforschung“, muss sich dafür aber bei den konkreten Äquivalenzproblemen manchmal mit allgemeinen Bemerkungen und Hinweisen zufriedustellen.

Die versprochene „praxisorientierte Einführung“ in die vergleichende Kulturforschung ist insbesondere in den beiden letzten Kapiteln des Buches zu finden. Thema von Kapitel 7 ist die „Erfassung sozialer Kontexte“ und ihr Vergleich auf der Basis von Input- oder Output-Harmonisierungsverfahren. Die ausführlichen Beispiele dieses Kapitels betreffen Klassifikationssysteme für soziodemographische Faktoren (Bildung, Einkommen und Beruf) und die Anwendung von Makroindikatoren (wirtschaftliche Entwicklung, politische Entwicklung). Die beiden Themen zeigen die Probleme und Möglichkeiten von Output-Harmonisierung an Hand von internationalen Standardisierungen in

bekannten Bereichen. Mit Harmonisierung wird zwar ein wichtiger Aspekt der Äquivalenzproblematik angesprochen. Es ist jedoch auch ein relativ einfacher Aspekt, der sich fast ausschließlich auf die Konstruktion und Validierung von eher identischen als von äquivalenten Instrumenten richtet.

Im letzten substanziellen Kapitel ihres Buches stellen Rippl und Seipel die Analyse-möglichkeiten von Daten vor, die in kulturvergleichenden Studien gewonnen werden. Dabei wenden sie einen systematischen Überblick basierend auf dem Unterschied zwischen Makrodaten, Makrodaten als aggregierte Werte von Mikrodaten und Mikrodaten an. Außerdem wird die Fehlschlussproblematik erläutert. Abschließend präsentieren die Autoren einige Beispiele von Äquivalenzprüfungen für eine „Nationalismus-Kurzskala im Kulturvergleich“. Auch hier wählen Rippl und Seipel leider nur relativ einfache Äquivalenzprobleme, da keines der Beispiele sich mit der Eliminierung oder Hinzufügung kulturspezifischer Items auseinandersetzt. Offensichtlich konnten die Autoren sich trotz allem nicht von dem Gedanken lösen, Identität oder Ähnlichkeit seien wesentliche Merkmale von Äquivalenz. Sie sind es nicht – und eine Auseinandersetzung mit der Konstruktion wirklich unterschiedlicher Instrumente in verschiedenen Kontexten fehlt leider komplett.

Mit diesem Buch haben Rippl und Seipel versucht, eine Lücke in dem Angebot von Lehrbüchern in der vergleichenden Sozialforschung zu schließen. Der Versuch ist sehr lobenswert und das Ergebnis kann sich durchaus sehen lassen. Wie bereits erwähnt, ist es in einem einführenden Lehrbuch der vergleichenden Forschung kaum zu vermeiden, auch allgemeine Aspekte der empirischen Sozialforschung zu behandeln. Für Dozenten, die das Buch in ihren Kursen und Seminaren einsetzen, bleibt allerdings noch viel zu tun. Erstens ist die ‚Praxisorientierung‘ des Buches auf Beispiele beschränkt, welche meistens nur knapp vorgestellt werden und relativ einfach sind. Die meisten Komplika-

tionen werden zwar erläutert, aber eine systematische Herangehensweise zur Reduzierung von Äquivalenzverletzungen fehlt. Hier hätte der Vorschlag von Fons van de Vijver und anderen, den Begriff ‚Äquivalenz‘ in kulturvergleichender Forschung durch ‚Verzerrung‘ (‚bias‘) zu ersetzen, sicherlich einen Ausweg darstellen können. Mit der Verwendung des Begriffes ‚Verzerrung‘ können verschiedene Ursachen von Vergleichbarkeitsproblemen und ihre Lösungen viel systematischer als mit dem unspezifischen Äquivalenzbegriff benannt werden.

Eine zweite Aufgabe, welche den Dozenten überlassen bleibt, betrifft die Anwendung von Ex-post Strategien in Sekundäranalysen. In den letzten Jahren ist das Angebot von vergleichenden Datenerhebungen (Eurobarometer, WVS, ESS, ISSP usw.) stark gestiegen und damit auch das Interesse an der Entwicklung von vergleichbaren Instrumenten. Manche dieser Erhebungen sind wiederholte Querschnittsbefragungen und somit auch auf die Entwicklung äquivalenter Messungen im Zeitvergleich angewiesen. Im Hinblick auf dieses stark gestiegene Datenangebot spielen Sekundäranalysen im Alltag der meisten vergleichenden Kulturforscher heutzutage eine zentrale Rolle. Auch hier bieten Rippl und Seipel leider wenig konkrete Hilfe. Ihre Darstellungen beschränken sich auf Input- und Output-Harmonisierungsverfahren und vermeiden den Unterschied zwischen Ex-ante und Ex-post Strategien, obwohl verschiedene Beispiele eindeutig zu dieser letzten Kategorie gehören.

Schließlich bleiben den Dozenten an mehreren Stellen auch Konkretisierungen des Informationsangebots überlassen. Von einem ‚praxisorientierten‘ Lehrbuch erwartet man zum Beispiel bei der Präsentation von Mehrebenenmodellen Hinweise auf die erforderlichen Fallzahlen für solche Analysen oder bei Makroanalysen Vorgaben für die Anwendung von Gewichtungen. Solche praktischen Informationen fehlen. Im Jahr 2008 sollte ein Buch für Studierende außerdem die Möglichkeiten des Internets viel breiter nutzen als die wenigen vereinzelten

Angaben von Websites, die Ripp und Seipel jetzt anbieten. Wer eine kompakte Einführung in die vergleichende Kulturforschung sucht, dem ist mit dieser Einführung sicherlich gut geholfen. Die Beschränkungen des Buches sollten dabei als Herausforderungen zu weiteren Explorationen des Gebietes betrachtet werden, wobei noch viele Probleme zu klären sind. Esping-Andersen lässt grüßen.

JAN VAN DETH, MANNHEIM

* * * * *



ROLF PORST,
2008: Fragebogen.
Ein Arbeitsbuch.
Studienskripten zur
Soziologie. Wiesba-
den: VS Verlag für
Sozialwissenschaften. ISBN: 978-3-
531-15178-6, 190
Seiten, 14,90 EUR.

Wer schon einmal einen Fragebogen entwickelt hat, der weiß, wie schwierig es ist, die richtigen Wörter zu finden, die Sätze so zu formulieren, dass sie für die Zielpersonen verständlich sind und den Fragebogaufbau so zu gestalten, dass er (in der mündlichen Befragung) vom Interviewer leicht zu handhaben ist bzw. dass er (in der schriftlichen Befragung) so übersichtlich aufgebaut ist, dass jede Zielperson ihn problemlos ausfüllen kann. Rolf Porst berät seit über 20 Jahren als wissenschaftlicher Mitarbeiter bei GESIS Forscher und Forschergruppen bei der Entwicklung und Durchführung von empirischen Studien und hat in dieser Zeit nach eigenen Angaben – „über den Daumen gepeilt“ (S. 5) – bei etwa 2.500 Fragebögen

Beratungshilfen gegeben. Es gibt vermutlich kaum einen Fehler, der bei der Konstruktion gemacht werden kann und den der Autor im Laufe dieser Zeit nicht gesehen hat. Das vorliegende Buch ist als Praxisbuch geschrieben, es soll sowohl dem empirisch arbeitenden Sozialwissenschaftler als auch dem Praktiker in der Marktforschung helfen, einen eigenen Fragebogen zu entwickeln. Der Autor verwendet dabei ganz bewusst eine Sprache, die dem Leser das Gefühl vermittelt, er säße gleich neben ihm. Er verwendet überwiegend Umgangssprache und spricht dabei den Leser immer wieder direkt an („Sie ahnen schon ...; sondern wir müssen die ...; wenn Sie dieses Zitat ...; wenn es mir gelungen sein sollte, Sie zu täuschen ...“ (S. 11)), um genau diese Situation zu simulieren – und er verwendet sie so, dass sie nicht überheblich klingt. Diese Art der Schreibweise ist im wissenschaftlichen Kontext etwas gewöhnungsbedürftig, aber nach zwei Kapiteln hatte ich mich daran gewöhnt und fand es anschließend sogar recht angenehm – man muss nicht darüber nachdenken, was der Autor meint oder meinen könnte, man nimmt seine Ratschläge zur Kenntnis. Die Arbeit hat auch einen gewissen Witz, u. a. wenn der Autor Beispiele für Fragenformulierungen bringt, die eigentlich absurd sind (aber sehr wahrscheinlich real waren).

Wie Porst ausführlich dokumentiert, hat jede Frage eine Antwort – und zwar jene, die zu dieser Frage passt. Es ist eine eigene Wissenschaft, es ist eine Kunst (richtig) zu fragen, oder wie es bereits Stanley Payne 1951 als Buchtitel formulierte „The art of asking questions“ bzw., gut 50 Jahre später in Anlehnung an diese Arbeit, von Schaeffer und Presser (2003) „The science of asking questions“. Wenn bedacht wird, dass bereits Payne (1951) sehr viele Hinweise auf Frageformulierungen und mögliche Fehler gab (und dieses Wissen wird auch in den Einführungen zu den Methoden der empirischen Sozialforschung gelehrt, es steht auch in jeder mir bekannten Einführung zu dem Thema) und dass es inzwischen jede Menge

weitere Erkenntnisse in dieser Richtung gibt (man denke u. a. an die Arbeiten in *Public Opinion Quarterly* und *International Journal of Public Opinion Research*), so ist es eigentlich verwunderlich, dass diese Fehler auch heute noch immer gemacht werden. Ich selbst könnte zahlreiche Beispiele geben, die sich mit denen des Autors (vgl. insbesondere S. 95–115) problemlos vertauschen ließen; ich denke dabei u. a. an so bekannte Studien wie das ISSP (z. B. die Fragen zum „political efficacy and trust“ im ISSP 2004 – und vergleichbare Fragen in vielen nationalen und internationalen Wahlstudien), wo es u. a. ‚negative Frageformulierungen‘ gibt, die bei Ablehnung eine doppelte Verneinung bedeuten (die aber aus guten Gründen – eigentlich – vermieden werden sollten). Doch, um es mit den Worten des Autors zu sagen, lassen Sie uns am Anfang des Buches beginnen.

Nach einer sehr motivierenden Einleitung folgen einige zentrale „kognitionspsychologische und kommunikative Grundlagen der Befragung“. Dabei wird deutlich, dass Porst nicht nur eine sehr praxisorientierte Einleitung schreibt, er zitiert hier (und ebenso in den folgenden Kapiteln) auch die relevante internationale Literatur zur Frageformulierung und zum Fragebogaufbau. Damit erhält der Leser neben Tipps für die eigene empirische Erhebung einen Überblick über zentrale Forschungsergebnisse, die meistens auf der Basis experimenteller Designs gewonnen wurden. Die Beispiele, die er in diesem und in den folgenden Kapiteln verwendet, sind zum einen aus der Literatur übernommen, zum anderen sind es eigene Beispiele aus Studien, an denen er im Rahmen der Umfragen von GESIS mitgearbeitet hat bzw. aus Projekten, bei denen er beratend tätig war.

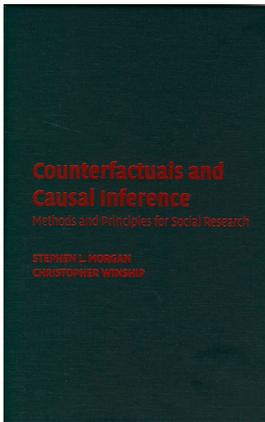
Die Ordnung der folgenden Kapitel entspricht soweit wie möglich dem Aufbau eines Fragebogens: Es beginnt mit einem Kapitel zur „Titelseite“, dann folgen „Hinweise zum Ausfüllen des Fragebogens“, es geht weiter mit „Arten von Fragen“ und „Arten von Skalen“ und endet mit „die ‚letzte Seite‘“,

dem „Layout des Fragebogens“ und „Befragungshilfen“; zu letzteren gehören z. B. Listen, anhand derer die Befragten das für sie Zutreffende wählen können. Die Schwerpunkte seiner Darstellung liegen in der face-to-face und schriftlichen Befragung, Besonderheiten der telefonischen Befragung werden nur relativ kurz abgehandelt. Nahezu vollkommen fehlen Hinweise auf Besonderheiten bei Online-Befragungen, hier hätte der Autor noch ein Kapitel (oder entsprechende Abschnitte in den Kapiteln) über die damit verbundenen Möglichkeiten einfügen können, z. B. über den Einsatz von virtuellen Interviewern und die Einbindung von Videos.

Wird das Buch insgesamt betrachtet, so legt Porst ein sehr praxisorientiertes Buch vor, in dem er sehr detaillierte Hinweise für die Konstruktion eines Fragebogens gibt. Er wendet sich sprachlich direkt an den Leser, indem er die persönliche Anrede und viele umgangssprachliche Formulierungen wählt. Dieser Schreibstil ist zwar eher ungewöhnlich, aber in einem ‚Praxisbuch‘ kann er anscheinend durchaus erfolgreich verwendet werden – und vielleicht ist er auch genau richtig, um auch ausgewiesene Forscher darauf hinzuweisen, dass auch sie sich an die Regeln der Frageformulierung halten sollten. Die Einführung wird abgerundet durch viele Hinweise auf die (internationale) Fachliteratur, so dass das Buch auch für diejenigen Leser nützlich ist, die sich über die Effekte von experimentellen Designs bei der Frageformulierung und der Fragenanordnung informieren möchten. Mir hat das Buch sehr gut gefallen und ich kann es sowohl Studierenden als auch Lehrenden nur empfehlen.

Literatur

- Payne, S., 1951: *The Art of Asking Questions*. Princeton: University Press.
- Schaffer, N. C. und S. Presser, 2003: *The Science of Asking Questions*. *Annual Review of Sociology*, 29, 65–88.



STEPHEN L. MORGAN und CHRISTOPHER WINSHIP, 2007: Counterfactuals and Causal Inference. Methods and Principles for Social Research. Cambridge: Cambridge University Press. ISBN: 978-0-512-85615-7 ISBN: 978-0-512-67193-4, 319+xiii Seiten, 23,99 EUR.

„Avoiding causal language when causality is the real subject of our investigation either renders the research irrelevant or permits it to be undisciplined by the rules of scientific inference. ... Rather we should draw causal inferences where they seem appropriate but also provide the reader with the best and most honest estimate of the uncertainty of that inference.“
(King G., R. O. Keohane und S. Verba, 1994: *Designing Social Inquiry*. Princeton, S. 76)

Es gibt wohl kaum eine Frage, die so grundsätzlich an das Wissenschaftsverständnis der empirischen Sozialwissenschaften rührt, wie jene nach den Bedingungen oder gar der prinzipiellen Möglichkeit adäquater kausaler Schlussfolgerungen in Bezug auf die Ursachen sozialen Handelns. Dies gilt insbesondere für die empirische Sozialforschung in der Soziologie und ihrer Nachbardisziplinen, die überwiegend nichtexperimentell arbeiten und deshalb in ihren Forschungsdesigns in der Regel nicht auf die experimentelle Manipulation des kausal relevanten Faktors einschließlich der Konstanthaltung sonstiger Einflussfaktoren mittels randomisiertem Laborexperiment zurückgreifen können. Dementsprechend legt bereits ein kurzer Blick in die Geschichte der empirischen Sozialforschung beredetes Zeugnis ab über das stete Auf und Ab der methodologischen Debatte um den kausalen Gehalt empirischer Befunde, die fortwährende Entwicklung von Techniken der Kausalanalyse sowie die oft enttäuschten Hoffnungen auf eine klare statistische Lösung der fundamentalen Problematik kausalen Schließens auf der Basis von Beobachtungsdaten.

In der Tat rollt seit etwa zehn Jahren eine weitere Welle der intensiven Beschäftigung mit der Methodologie kausalen Schließens durch die Sozialwissenschaften. Insbesondere die Arbeiten von Donald Rubin, Paul Rosenbaum, James Robins oder Judea Pearl in der Statistik sowie von James Heckman, Charles Manski oder Joshua Angrist in der Ökonometrie haben dazu beigetragen, die methodologischen Bedingungen valider Kausalschlüsse konzeptionell klar zu fassen und dadurch nicht zuletzt für die angewandte empirische Sozialforschung den Zusammenhang zwischen Forschungsdesign und adäquater statistischer Modellspezifikation in der Analyse kausaler Effekte zu verdeutlichen. Verglichen mit den intensiven Debatten in Statistik, Ökonometrie und jüngst auch verstärkt in der Politikwissenschaft hat die Diskussion um das so genannte kontrafaktische Kausalitätsverständnis in der Soziologie abseits der methodologisch spezialisierten Journale nur vergleichsweise wenig Widerhall in der Methodendiskussion oder gar der Forschungspraxis gefunden. Und soweit die soziologische Aufmerksamkeit geweckt werden konnte, gebührt nicht zuletzt Stephen Morgan und Christopher Winship ein gut Teil des Verdienstes, als Autoren ihres bekannten Überblicksartikels im Jahrgang 1999 des *Annual Review of Sociology* die statistische und ökonometrische Literatur für die Soziologie zugänglich gemacht zu haben.

Counterfactuals and Causal Inference. Methods and Principles for Social Research ist nun letztlich die logische Fortsetzung des Programms der beiden amerikanischen Wissenschaftler. Ganz im Einklang mit der programmatischen Eingangsformulierung aus King, Keohane und Verbas einschlägigem Lehrbuch geht es Morgan und Winship um nichts weniger als die Klärung der Bedingungen und Schwierigkeiten kausaler Inferenz in der empirischen Sozialforschung. Mittel zum Zweck ist dabei die umfassende Darstellung des kontrafaktischen Kausalitätsverständnisses als solide Grundlage der

Kausalanalyse in den nichtexperimentellen Sozialwissenschaften, sowie die Diskussion der sich aus dem kontrafaktischen Ansatz ergebenden Implikationen für die statistische Analyse von Beobachtungsdaten. Im ersten Teil des Buches steht daher zunächst die vertiefte Diskussion des kontrafaktischen Modells im Vordergrund, wobei Morgan und Winship zentral auf die Arbeiten Judea Pearl's rekurrieren, um zu einer einheitlichen Präsentation zu gelangen. Für die Rezeption des Modells in der soziologischen Praxis dürfte diese Entscheidung tatsächlich durchaus hilfreich sein, da dadurch an eine Darstellung in Form des weit verbreiteten pfadanalytischen Kausaldiagramms angeknüpft werden kann.

In ihrer Darstellung unterstreichen Morgan und Winship zu Recht, dass sich der kontrafaktische Ansatz eben nicht in der Proliferation neuer Terminologien wie etwa des durchschnittlichen kausalen Effekts (average treatment effect) und ähnlicher Wortungetüme erschöpft, sondern dass die zentrale Leistung des Modells darin besteht, Fragen der empirischen Identifikation kausaler Effekte wieder in den Mittelpunkt der methodischen Debatte in der empirischen Sozialforschung zu rücken. Mit dem kontrafaktischen Modell gelangt die experimentelle Terminologie also zurück in die sozialwissenschaftliche Analyse von Beobachtungsdaten, und zwar „as a framework in which to ask carefully constructed ‚what-if‘ questions that lay bare the limitations of observational data and ... that can help researchers to determine what assumptions must be maintained in order to identify causal effects, as well as the selection of an appropriate technique that can be used to estimate an identified causal effect from a sample of sufficient size under whatever assumptions are justified“ (S. 13 und 23).

Mit der Betonung der Identifikationsproblematik durch Morgan und Winship wird deutlich, dass im kontrafaktischen Verständnis dem Zusammenspiel von theoretischer und statistischer Analyse die ent-

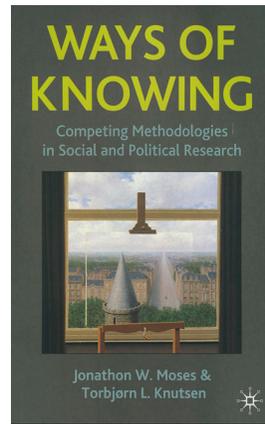
scheidende Bedeutung in der Ermittlung kausaler Effekte in der empirischen Sozialforschung zukommt. Anders als in mancher Spielart experimenteller Forschung ist eine theoriefreie statistische Analyse von Beobachtungsdaten letztlich unfruchtbar, da die Identifizierbarkeit eines bestimmten kausalen Effekts sich immer aus dem seitens des Forschers unterstellten theoretischen Modells ergibt. Gleichzeitig zeigen Morgan und Winship etwa durch die Einbeziehung der klassischen Lazarsfeldschen Begrifflichkeiten der antezedierenden und intervenierenden Variablen in den Zusammenhang des kontrafaktischen Modells, dass inhaltlich-theoretischen Überlegungen selbstverständlich eine zentrale Rolle in der statistischen Modellspezifikation zukommt.

Insofern präsentiert *Counterfactuals and Causal Inference* im stärker angewandten zweiten und dritten Teil eben kein Sammelurium neuer und neuester statistischer Schätzalgorithmen, sondern diskutiert die Anwendung klassischer statistischer Techniken der empirischen Sozialforschung aus dem Blickwinkel des kontrafaktischen Modells und unter der Zielsetzung, kausale Effekte zu ermitteln. Teil 2 widmet sich dabei der Diskussion konditionaler Analysestrategien, wobei die Darstellung nichtparametrischer Matchingverfahren (Kapitel 4) und traditioneller Regressionsansätze (Kapitel 5) zentralen Raum einnimmt. Angesichts ihrer Bedeutung sei hier natürlich insbesondere das Kapitel zur kontrafaktisch interpretierten Regressionsanalyse zur Lektüre empfohlen, deren Zielsetzung im kontrafaktischen Modell eben die Analyse der „effects of causes“ und nicht länger die traditionelle Betrachtung einer langen Reihe potenziell konkurrierender „causes of effects“ darstellt. Ebenso wichtig scheint Morgan und Winships Hinweis auf die strukturelle Verwandtschaft von Regressions- und Matchingverfahren in Bezug auf die zugrunde liegende kausalanalytische Strategie – womit hoffentlich einer neuen, vorhersehbar enttäuschten Hoffnung auf das statistische Allheilmittel des Propensity

Score Matching vorgebeugt sei. In Teil 3 werden schließlich alternative und in der Soziologie bislang noch (zu) wenig verbreitete Verfahren vorgestellt – insbesondere die Instrumentalvariablen-schätzung (Kapitel 7), die Analyse von Längsschnittdaten mittels Panelregressionsmodellen (Kapitel 9) und Pearls Vorschlag der Identifikation kausaler Effekte über isolierte Mechanismen (Kapitel 8), die gewinnbringend in der Analyse kausaler Wirkungen eingesetzt werden können, falls über ein Querschnittformat hinausgehende Datenstrukturen vorliegen. Angesichts der Fülle des Materials und der beispielhaften Stringenz der Darstellung, die Morgan und Winship hier gelingt, besteht kein Zweifel, dass mit Counterfactuals and Causal Inference die bislang umfassendste Einführung in das kontrafaktische Kausalitätsverständnis und seine Konsequenzen für die empirische Sozialforschung vorliegt. Morgan und Winship leisten damit einmal mehr einen Beitrag dazu, die Soziologie nachhaltig am aktuellen Stand der methodologischen Diskussion in ihren Nachbar-disziplinen zu orientieren. Es wäre dementsprechend zu wünschen, dass sich dieses Buch zur Standardlektüre in den Graduiertenschulen und der Praxis der empirischen Sozialforschung entwickelt.

MARKUS GANGL, MADISON

* * * * *



JONATHON W. MOSES und TORBJØRN L. KNUTSEN, 2007: *Ways of Knowing. Competing Methodologies in Social and Political Research.* Palgrave Macmillan. ISBN 10: 978-0-230-51665-6, 344 Seiten, 33,99 EUR.

„Ways of Knowing“ ist in mancherlei Hinsicht ein bemerkenswertes Buch. Unter anderem auch aus dem Grund, dass es sich hierbei weniger um eine der mehr oder weniger üblichen Einführungen in die Methodologie der Sozialwissenschaften handelt, sondern in einem gewissen Sinne um eine Ideengeschichte verschiedener methodologischer Ansätze. Dies trifft insbesondere für den ersten Teil zu. Insofern ist das Buch auch weniger eine Methodologie im engeren Sinn, auch kein Buch über die diversen Methodologien, sondern vielmehr in großen Teilen ein Werk, das die wissenschaftsphilosophischen Hintergründe der jeweiligen Methodologien aufzeigen will. Aufgrund dieser Einordnung der Methodologien in bestimmte erkenntnis- und wissenschaftstheoretische Traditionen stellt das Buch von Moses und Knutsen eine informative Lektüre dar, die dem Leser hilft, sich in der gegenwärtig geführten Diskussion und oft streitbaren Auseinandersetzung zwischen so genannten ‚quantitativen‘ und ‚qualitativen‘ Methodologien zurecht zu finden. Die Autoren selbst vertreten einen pluralistischen Ansatz, indem sie dazu auffordern, die verschiedenen Methoden bzw. Methodologien wie eine Werkzeugkiste zu betrachten, aus der man sich die jeweils passende Methode für das entsprechende Problem sucht.

Die beiden erwähnten methodologischen Traditionen werden von Moses und Knutsen als die naturalistische und die konstruktivistische Sichtweise bezeichnet, da sie das Begriffspaar ‚quantitativ/qualitativ‘ mit guten Gründen als irreführend zurückweisen. Nach der Einleitung widmen sich die folgenden sechs Kapitel der naturalistischen Vorgehensweise. Im zweiten Kapitel wird die Philosophie naturalistischer Wissenschaft erläutert. Hierbei greifen die Autoren auf die üblichen Väter der modernen Wissenschafts- und Erkenntnistheorie wie Bacon, Locke, Hume und Descartes zurück. Die neuere Wissenschaftsphilosophie wird mit dem Wiener Kreis, Popper und der Post-Popper-Tradition, hier vor allem Hempel, abgehandelt. Die ‚Hierarchie‘ naturalistischer Methoden von Experiment, Statistik und Vergleich wird in den folgenden drei Kapiteln behandelt. Danach schließt sich ein Kapitel über ‚Geschichte und Fallstudien‘ an. Das Erfrischende an der Erörterung dieser Methoden ist, dass sie weniger in der gegenwärtig üblichen Form präsentiert werden, sondern in der paradigmatischen Form der klassischen Arbeiten zu den jeweiligen Methoden. So ist zwar Mill auch sonst ein fester Bestandteil jedes Methodenbuchs zur vergleichenden Forschung, ungewöhnlich aber ist der Verweis auf Graunt, Petty und Galton, um die Logik deskriptiver und inferenzieller statistischer Verfahren zu erläutern, sowie die relativ ausführliche Behandlung der historischen Methode von Leopold von Ranke. Lijphart und Eckstein als Referenzautoren für die Fallstudie nehmen dagegen wieder sehr konventionelle Züge an, wobei man hier womöglich auch neuere Autoren wie z. B. Gerring oder George und Bennett hätte aufnehmen können.

Das siebte Kapitel weckt Zweifel über die Angemessenheit der naturalistischen Methode im Bereich der Sozial- und Humanwissenschaften, also dann, wenn es bei den zu untersuchenden Gegenständen um Folgen von Willenshandlungen von Personen geht, und leitet so zur konstruktivistischen Sicht-

weise über. Die Philosophie der konstruktivistischen Sichtweise wird einerseits mit der Erkenntnistheorie Kants in Abgrenzung zu Hume erläutert, im Wesentlichen aber wieder durch Referenz auf einen anderen inzwischen weitgehend Unbekannten, nämlich William Whewell. Das nächste Kapitel stellt eher wieder auf konventionelle Art die konstruktivistische Methode vor, es geht um das Erzählen von Geschichten und Geschichte. Während dies in klarer Abgrenzung zur naturalistischen Methode steht, d. h. sich vor allem unter dem Hinweis auf die dort vorliegende Vernachlässigung der Bedeutung des Kontexts vollzieht, widmen sich die folgenden drei Kapitel gewissermaßen den Ähnlichkeiten zwischen naturalistischer und konstruktivistischer Sichtweise, indem sie darauf hinweisen, dass die paradigmatischen Methoden der Naturalisten, Experiment, Statistik und Vergleich, auch unter dem konstruktivistischen Vorzeichen Verwendung finden bzw. Verwendung finden könnten. Damit die spiegelbildliche Beziehung erkennbar wird, werden die drei Methoden in umgekehrter Reihenfolge erläutert wie innerhalb des Rahmens der naturalistischen Philosophie. Dabei entspricht allerdings das erste Kapitel über den interpretierenden Vergleich, bzw. genauer die vergleichenden Interpretationen, noch weitgehend dem Standard konstruktivistischer Methodologie. Als innovativ und im höchsten Maße Gedanken (und Widerspruch) provozierend können hingegen die beiden darauf folgenden Kapitel bezeichnet werden. Im elften Kapitel zeigen die Autoren drei Methoden auf, wie ihrer Meinung nach der Kontext stärker bei statistischen Verfahren berücksichtigt werden könnte, womit die Statistik wohl konstruktivistisch gewendet werden soll. Diese drei Methoden bestehen aus der Anwendung komplexer grafischer Illustrationen, der bayesianischen Statistik und so genannter *Statistic(al) Narratives*, in Anlehnung an die konstruktivistische Wendung des *Rational Choice Ansatzes* durch die *Analytical Narratives* von Bates et al. Hier scheinen mir allerdings

Zweifel angebracht, ob diese optimistische Sichtweise gerechtfertigt ist, denn der spontane Eindruck über den derzeitigen Boom von bayesianischen Methoden in der empirischen Forschung ist doch eher der, dass es sich bei deren Vertretern nicht gerade um diejenigen der naturalistischen Zunft handelt, die den erkenntnistheoretischen Prämissen des Konstruktivismus besonders offen gegenüberstehen. Das zwölfte Kapitel schließt die Klammer der Diskussion durch die Behandlung von Gedankenexperimenten als dem konstruktivistischen Äquivalent zu der klassischen Form des naturalistischen Experiments. Auch hier scheinen mir die Autoren, möglicherweise einem zu starkem Drang folgend, den spiegelbildlichen Aufbau zu erhalten, etwas über das Ziel hinauszuschießen. Die angeführten Gedankenexperimente rekurrieren meiner Meinung nach eher auf das Gegensatzpaar ‚normativ/‘positiv‘ als ‚konstruktivistisch/‘naturalistisch‘.

"Ways of Knowing" stellt einen – im besten Sinne des Wortes – umfangreich gebildeten Streifzug durch die philosophischen Traditionen dar, auf denen die naturalistische und die konstruktivistische Methodologie fußen. Selbst wenn man nicht mit allen Argumenten der Autoren konform gehen mag, so vermittelt das Buch in jedem Fall eine Fülle von äußerst interessanten wissenschaftsphilosophischen Argumenten und wissenschaftshistorischen Fakten, so dass die Lektüre allemal lohnend ist.

Die gegenwärtige Diskussion zwischen Vertretern der naturalistischen und konstruktivistischen Sichtweise ist oft nicht frei von Borniertheit und Arroganz. So kann man der methodologischen Propaganda des wohl derzeit einflussreichsten Vertreters des naturalistischen Lagers, Gary King, imperialistische Züge nicht immer ganz abstreiten, und die großzügige Art in „Designing Social Inquiry“, bestimmte erkenntnistheoretische Probleme zu lösen, indem man sie schlicht ignoriert, kann einem einerseits als im Streit um methodologische Deutungshoheit aus

strategischen Gründen praktizierte Chuzpe Bewunderung abringen, hat aber andererseits nicht wenige Wissenschaftler irritiert, die nicht grundsätzlich die Welt aus Nägeln bestehend wahrnehmen wollen, nur weil sie über den entsprechenden (statistischen) Hammer verfügen. Umgekehrt sehen sich statistisch und empirisch arbeitende Sozialwissenschaftler von Seiten des konstruktivistischen Lagers nicht selten dem Generalverdacht ausgesetzt, dass die mangelnde Berücksichtigung der Komplexität der Welt in ihren Modellierungen letztlich nur den eigenen persönlichen Mängeln an kognitiver Komplexität geschuldet ist, die sie zu einem unreflektierten Dasein als positivistische Erbsenzähler verdammen. Das Buch von Moses und Knutsen kann möglicherweise seinen Teil dazu beitragen, diese bestehenden Gräben zuzuschütten. Allerdings sind sie nicht die einzigen, die an einer solchen versöhnlichen bzw. integrierenden Sichtweise arbeiten, insofern ist es verwunderlich, dass solche Ansätze, die oft unter dem Begriff ‚mixed methodology‘ gefasst werden, in dem Buch keine Beachtung finden. Es existieren lediglich vage Andeutungen auf diese von den Autoren "realistisch" genannte Sichtweise. Auch bin ich mir nicht sicher, ob die Autoren überhaupt eine pluralistische Sichtweise vertreten, wie von ihnen des Öfteren selbst behauptet. Zumindest sind sie es ganz eindeutig nicht im Sinne von Feyerabends "Anything goes", in diesem Sinne aber ist meiner Meinung nach dieser Begriff häufig besetzt. Es geht daher wohl weniger um Methodenpluralismus als um Methodenvielfalt. Um beim Bild der Werkzeugkiste zu bleiben: Für manche Probleme braucht man einen Hammer, für andere einen Schraubendreher. Wer jedoch für Probleme, die einen Hammer benötigen, einen Schraubendreher einsetzt, ist ein schlechter Handwerker und im übertragenen Sinne eben ein schlechter Wissenschaftler. Jeder gute Handwerker aber handelt im Bewusstsein und in Kenntnis der Tradition seines Handwerks, selbst wenn es sich dabei um Tipps und Kniffe handeln

sollte, die nicht mehr angewandt werden. Dies scheint mir daher das größte Verdienst dieses Buches zu sein: dem empirisch Forschenden die Augen zu öffnen für die Tradition, in der er jeweilig steht bzw. diejenigen, die sich anderer Methoden bedienen. Auch wenn hier schon vieles gesagt worden ist, leisten Moses und Knutsen wertvolle Aufklärungsarbeit in dieser Hinsicht. Es würde nicht zum Schaden der Wissenschaft sein, wenn dieses Buch weit reichende Aufmerksamkeit erfährt, sein Einsatz in einführenden Seminaren zu diesem Thema wird jedenfalls lohnend für die Studierenden sein.

JOACHIM BEHNKE, FRIEDRICHSHAFEN

Ankündigungen

Mobile Research Conference line up announced: Dr. Liz Nelson to Keynote

*Globalpark, London
February 16 - 17, 2009*

Mobile Research Conference

The organizing committee of the Mobile Research Conference is delighted to announce that Dr. Liz Nelson, Chairman of Q-Research, will provide a keynote speech at the event in London next February. Liz will be discussing the potential of mobile research and implications for industry standards.

In the market research industry the name of Dr. Liz Nelson is legendary. As founder and chairman of Taylor Nelson Sofres she helped develop the market research principles that have formed the foundation of today's industry. TNS plc was the second largest market research company in the world at the time of its recent acquisition by WPP.

Tim Macer, founder of Meaning Ltd, a consultancy specializing in research technology will be speaking on mobile technology in research. With over 25 years experience as a research technology practitioner, Tim is respected as an independent and authoritative observer and advocate for the application of software and informatics in survey research.

The conference is hosted by Globalpark, a leading provider of online feedback software and mobile research technology. Delegates will discuss the opportunities and challenges presented by the evolution of the mobile as a research platform. Over a day and a half they will hear the experiences, perspectives and techniques of market researchers from both corporate and academic backgrounds.

Scientific and academic pioneers of this emerging research approach will also be participating, including Prof. Dr. Michael Bosnjak of the Free University of Bozen-Bolzano, Prof. Dr. Marek Fuchs of The University of Kassel, and Prof. Dr. Vasja Vehovar of The University of Ljubljana. We are delighted to have speakers from leading global market research companies including Ipsos, Lightspeed Research, Common Knowledge, The 3rd Degree, The Nielsen Company and TNS Gallup.

The conference takes place in London, February 16-17, 2009. Further details and registration is available at www.mobileresearch09.com.

**European Labour Force Survey
(EU-LFS) and
European Union Statistics on Income
and Living Conditions (EU-SILC):
1st European User Conference**

Organized by
German Microdata Lab, GESIS,
in cooperation with Eurostat

Mannheim

March 5 - 6, 2009

*European User
Conference
EU-LFS and
EU-SILC*

The first European User Conference for EU-LFS and EU-SILC, which is organized by the German Microdata Lab, GESIS, in cooperation with Eurostat, will provide researchers who use these data with the opportunity to discuss their work and share their experience. In addition to fostering the discussion within the research community on both substantive and methodological issues, the conference offers researchers the opportunity to give feedback to the European Statistical System. It is important that the potential and limitations of EU-LFS and EU-SILC are understood by researchers and, at the same time, that the needs of researchers become clear.

Researchers of all disciplines who use either EU-LFS or EU-SILC micro data are encouraged to participate. Conference sessions will focus on education, labour market, income, poverty and methodological topics. The detailed program can be found on the next pages and on the conference webpage at <http://www.gesis.org/EU-User-Conference/>.

REGISTRATIONS AND CONFERENCE FEE:

Registrations can be made via eu-microdata@gesis.org. The conference fee is 120 Euros for participants (students: 90 Euros). The fee covers coffee breaks, conference dinner and lunch.

Local organizers of the conference are Christof Wolf (christof.wolf@gesis.org) and Heike Wirth (heike.wirth@gesis.org). For any further questions please contact the local organizers and visit the conference webpage: <http://www.gesis.org/EU-User-Conference/>.

European Labour Force Survey and European Union Statistics on Income and Living Conditions: 1st European User Conference – Program

Thursday, March 5, 2009

9.00	Registration		
10.00-11.30	Opening session		
11.30-12.30	Plenary session		
	Christopher Whelan and Bertrand Maître Comparing poverty indicators in an enlarged EU		
	Michael Gebel and Johannes Giesecke Labour market flexibility and inequality: the changing skill-based temporary employment and unemployment risks in Europe		
12.30-14.00	Lunch		
14.00-17.30	Session A: Labour Market	Session B: Poverty	Session C: Methods
	Andreas Ette & Leonore Sauer Measuring spatial mobility with the European Union Labour Force Survey: The case of intra-European migrants	Silvia Avram Impact of social assistance schemes in Central and Eastern Europe	Claudio Ceccarelli & Lucia Coppola Selective attrition in EU-SILC: evidence from the first three waves
	Alberto Veira Ramos The role of occupational and socio-demographic determinants on the age of retirement in Spain	Paolo Consolini and Gabriella Donatiello Individual poverty risk and household income patterns: A comparison at European Level	Katarzyn Saczuk & Pawel Strzelecki The impact of migration on the LFS data: The case of Poland
	Leila Maron & Danièle Meulders The parenthood effect on employment in Europe	Henning Lohmann Measuring poverty and employment using EU-SILC: Differences between register and survey data	Ann-Catherine Guio, Alessio Fusco & Eric Marlier Material Deprivation in the European Union: results and issues drawing from EU-SILC and Eurobarometer data
15.30-16.00	Coffee		
	Marie Plessz Using EU-LFS data for cohort analysis in Central Europe: inter-cohort inequality in labour market outcomes	Meulders, Danièle, Kim Fredericq Evangelista & Zouhair Alaoui Amine Individualisation of income and poverty risk: a comparative analysis in 24 European countries	Matthias Till, Johannes Klotz & Bernhard Meindl Enlarging small domains: an application of EU-SILC regression parameters to the LFS for obtaining disaggregated poverty estimates
	Alessandra Pescarolo Work quality: comparing 6 European regions	Eva Sierminska The economic vulnerability of the elderly in a cross-national perspective: exploring the EU-SILC	Luis Toharrio & Virginia Hernanz The unemployment and inactivity in UK, France and Spain: The treatment of overlap areas (grey areas).
	Antonio R. Disceenza, Barbara Boschetto, Francesca Fiori, Carlo Lucarelli, Simona Rosati Occupational transitions in Italy from 2004 to 2008	Maria Noel Pi Alperin & Philippe van Kerm Dynamics of multidimensional poverty in EU countries	Ralf Münnich, Jan-Philipp Kolb & Stefan Zins Accuracy of poverty indicators derived from EU-SILC
19.00	Dinner		

Friday, March 6, 2009

9.00-11.45	<p>Session D: Education Walter Müller & Markus Klein Bildungsgleichheit im internationalen Vergleich</p> <p>Sophie Dantan Credit constraints on educational attainment: evidence from modelling a discrete choice model with latent choice sets</p> <p>August Gächter, Dilek Çınar, Holger Seibert, Drenka Vukovic & Natalija Perisic How well does education travel? Education and occupation with and without migration</p> <p>Break</p> <p>David Reimer, Clemens Noeike & Aleksander Kucel Labor market effects of field of study in comparative perspective</p> <p>Stan van Alphen The benefits of educational expansion for early school leavers on the European labour market</p> <p>Coffee</p> <p>Plenary Session</p> <p>Carmen Petrovici & Ruud J.A. Muffels Welfare state institutions and early retirement in Europe: a comparative analysis using EU-SILC</p> <p>Closing Discussion</p>	<p>Session E: Gender Gap Alena Bicakova Gender unemployment gaps: evidence from the new member States</p> <p>Yekaterina Chzen Decomposing gender wage gaps across the distribution in Europe: including allowance for sample selection adjustment</p> <p>Martin Kubiček & Martina Mysiková Income inequalities within couples in the Czech Republic and selected European countries</p> <p>Sile O'Dorchai Do women gain or lose from becoming mothers? A comparative wage analysis in 25 European countries</p> <p>Catia Nicodemo & Robert Waldmann Child-care and participation in the labour market for married women in Mediterranean countries</p>	<p>Session F: Income Richard Heuberger Housing and social exclusion: considering imputed rents, interest repayments and housing costs for the household income in EU-SILC</p> <p>Richard Berthoud, Francesco Figari, Maria Iacovou & Holly Sutherland Income distribution and redistribution across Europe: using EU-SILC in two complementary approaches</p> <p>Judith Niehues Income inequality within and between European countries</p> <p>Heinz-Herbert Noll & Stefan Weick Making ends meet? Income, standards of living and subjective assessments across European countries</p> <p>Jonas Radl Direct or sequential retirement? The selectivity of exit pathways in Denmark, Germany, and Spain</p>
10.30-10.45			
11.45-12.15			
12.15-13.30			

Call for Papers

6. Nutzerkonferenz „Forschung mit dem Mikrozensus“ Analysen zur Sozialstruktur und zum sozialen Wandel

GESIS, Mannheim

15. – 16. Oktober 2009

Analysen zur Sozialstruktur und zum sozialen Wandel

Die 6. Nutzerkonferenz widmet sich der Untersuchung der Sozialstruktur und des sozialen Wandels in Deutschland. Auf der Basis von Mikrozensusdaten gewonnene Forschungsergebnisse sollen vorgestellt und diskutiert werden. Darüber hinaus ist die Konferenz ein Forum für den Erfahrungsaustausch der Datenutzer/innen untereinander sowie mit den Vertreter/innen der amtlichen Statistik. Sie wendet sich an Wissenschaftler/innen, die mit den Scientific Use Files des Mikrozensus arbeiten oder zukünftig mit diesen Daten arbeiten wollen. Gegenwärtig wird mit den Scientific Use Files der Zeitraum 1973 bis 2006 abgedeckt. Zudem stehen für Analysen im individuellen Längsschnitt Paneldaten der Mikrozensus der Jahre 1996 bis 1999 sowie 2001 bis 2004 als Scientific Use Files zur Verfügung.

Es können Beiträge aus allen Arbeitsgebieten eingereicht werden, die sich unter Verwendung von Scientific Use Files des Mikrozensus mit Fragen der Sozialstruktur und des sozialen Wandels oder mit methodischen Aspekten bzw. der Datenqualität beschäftigen.

Interessent/innen werden gebeten, Abstracts der geplanten Beiträge bis spätestens 30. März 2009 einzureichen an: andrea.lengerer@gesis.org und julia.schroedter@gesis.org. Die Abstracts sollten maximal zwei Seiten umfassen sowie den Titel des Beitrags und Name, E-Mail und Institution der Autor/innen enthalten.

Durchgeführt und organisiert wird die Konferenz vom German Microdata Lab (GML) der GESIS und vom Statistischen Bundesamt, Gruppe VIIC – Mikrozensus.

Ihre Ansprechpartnerinnen bei GESIS sind:

Andrea Lengerer, Julia Schroedter

GESIS – Leibniz-Institut für Sozialwissenschaften

German Microdata Lab

Postfach 12 21 55

D - 68072 Mannheim

Tel.: +49 (0)621 1246 265; Fax: +49 (0)621 1246 100

andrea.lengerer@gesis.org; julia.schroedter@gesis.org

www.gesis.org/gml/veranstaltungen

Autorinnen und Autoren Jahrgang 2 (2008)

- Christian Babka von Gostomski, Nürnberg
- Dorothea Böhr, Leipzig
- Marc Bühlmann, Zürich
- Gregor Daschmann, Mainz
- Martin Kroh, Berlin
- Ivar Krumpal, Leipzig
- Elias Naumann, Leipzig
- Rainer Pischner, Berlin
- Andreas Pöge, Bielefeld
- Monika Pupeter, München
- Heiko Rauhut, Leipzig
- Sabine Schiller-Born, Berlin
- Patrick Schmich, Berlin
- Anna Schnauber, Mainz
- Martin Spieß, Hamburg
- Isabelle Stadelmann-Steffen, Konstanz
- Gert G. Wagner, Berlin
- Christian Weidmann, Berlin

Rezensentinnen und Rezensenten Jahrgang 2 (2008)

- Joachim Behnke, Friedrichshafen
- Jörg Blasius, Bonn
- Jan van Deth, Mannheim
- Hartmut Esser, Mannheim
- Marek Fuchs, Kassel
- Markus Gangl, Madison
- Peter Hartmann, Düsseldorf
- Anthony A. Oberschall, Chapel Hill
- Bernd Weiß, Köln

Gutachterinnen und Gutachter Jahrgang 2 (2008)

Wir danken folgenden Kolleginnen und Kollegen, die für den Jahrgang 2 (2008) der Methoden – Daten – Analysen Manuskripte begutachtet haben:

- Tobias Bachteler, Konstanz
- Jörg Blasius, Bonn
- Michael Blohm, Mannheim
- Claudia Diehl, Göttingen
- Detlev Fetchenhauer, Köln
- Frauke Kreuter, Maryland
- René Leicht, Mannheim
- Oliver Lipps, Lausanne
- Rainer Metz, Köln
- Jochen Musch, Düsseldorf
- Walter Müller, Mannheim
- Ulrich Pötter, Bochum
- Götz Rohwer, Bochum
- Kurt Salentin, Bielefeld
- Elmar Schlüter, Mannheim
- Rüdiger Schmitt-Beck, Mannheim
- Jörg-Peter Schräpler, Bochum
- Volker Stocké, Mannheim

Hinweise für unsere Autorinnen und Autoren

Methoden – Daten – Analysen (MDA) veröffentlicht Beiträge aus dem Bereich der Empirischen Sozialforschung, insbesondere aus dem Bereich der Umfragemethodik. Im Vordergrund stehen Artikel, welche die methodischen und/oder statistischen Kenntnisse der Profession erweitern, sowie Beiträge, die sich mit der Anwendung der Methoden der Empirischen Sozialforschung in der Forschungspraxis beschäftigen, oder solche, in denen ein statistisches Verfahren exemplarisch angewandt wird. Obwohl der Schwerpunkt auf Umfragemethoden liegt, sind Beiträge zu anderen methodischen Bereichen willkommen. Die Artikel sollen für eine breite Leserschaft von Wissenschaftlern und Praktikern im Bereich der Empirischen Sozialforschung verständlich sein.

Manuskripte, die bereits an anderer Stelle veröffentlicht sind oder gleichzeitig anderen Publikationsorganen zur Veröffentlichung angeboten worden sind, werden grundsätzlich nicht berücksichtigt. Eine spätere Veröffentlichung eines in der MDA erschienenen Beitrages ist möglich, sofern an exponierter Stelle auf die Ersterscheinung des Beitrages in der MDA hingewiesen wird.

Jeder Beitrag, der zur Veröffentlichung in MDA eingereicht wird, wird zunächst von den Herausgebern danach bewertet, ob er für eine Veröffentlichung grundsätzlich in Frage kommt.

Falls die Herausgeber einer Veröffentlichung grundsätzlich ablehnend gegenüber stehen, werden die Autoren unter Angabe von Gründen für diese Entscheidung informiert.

Falls die Herausgeber zur Ansicht gelangen, dass der Beitrag grundsätzlich zur Veröffentlichung in Frage kommt, wird er anonymisiert an mindestens zwei unabhängige Gutachter verschickt, die um eine Stellungnahme gebeten werden. Im Zweifelsfalle wird ein drittes Gutachten eingeholt.

Wird ein Beitrag nach Beschluss der Herausgeber in das Begutachtungsverfahren gegeben, erfolgt die abschließende Entscheidung über ein Manuskript auf der Basis der Gutachten durch die Herausgeber. Im Falle einer Ablehnung erhalten die Autoren eine ausführliche Begründung für die Ablehnung. Wird eine Überarbeitung eines Beitrages für erforderlich gehalten, erhalten die Autoren detaillierte Überarbeitungshinweise.

Unabhängig vom Ergebnis des Begutachtungsverfahrens werden die Autoren von der Entscheidung durch die Redaktion per E-Mail informiert.

Die folgenden Regeln sind bei der Abfassung von Manuskripten zu beachten:

Manuskripte müssen per E-Mail (mda@gesis.org) eingereicht werden. Der Umfang der Manuskripte soll inklusive Leerzeichen alles in allem nicht mehr als 70.000 Zeichen betragen.

Den Beiträgen sind Abstracts in Deutsch und Englisch (jeweils ca. 15 Zeilen) voranzustellen. Auch der Titel des Beitrages ist in Deutsch und Englisch einzureichen.

Um die Anonymität der Beiträge zu wahren, darf in einem Manuskript nur der Titel des Beitrages enthalten sein, nicht aber Namen oder Anschriften der Autoren; Name und Anschrift der Autoren müssen, gemeinsam mit dem Titel des Beitrages, auf einer separaten Seite eingereicht werden.

Beiträge sind mit dem Dezimalklassifikationssystem zu untergliedern (1 - 2 - 2.1 - 2.2 - 3 usw.). Die Gliederungstiefe geht dabei höchstens auf *eine* Stelle nach dem Punkt.

Tabellen enthalten Tabellennummer und Titel im Tabellenkopf, Abbildungen werden analog behandelt.

Grafiken sind mittels gängiger Grafiksoftware zu erstellen. Ist eine spezielle Grafiksoftware erforderlich, übernimmt der Autor/die Autorin die endgültige Formatierung der Grafiken in eigener Regie.

Bei der Erstellung von Tabellen und Grafiken ist zu berücksichtigen, dass der Satzspiegel 11,5 cm (Breite) x 18,5 cm (Höhe) beträgt. Die Grafiken sind als jpeg- oder tif-Dateien in schwarz-weiß mit einer Auflösung von mindestens 150 dpi zu liefern.

Die Beiträge sind unter Wahrung der gültigen Rechtschreiberegeln (neue Rechtschreibung) zu erstellen.

Werden in einem Beitrag empirische Daten verwendet, muss die Möglichkeit der Replikation bestehen. Im Falle einer Veröffentlichung in der MDA erklären sich die Autoren daher schriftlich bereit, Dritten auf deren Anfrage hin die Daten und ProgrammROUTINEN zur Verfügung zu stellen.

Anmerkungen und Fußnoten sind mit der Fußnotenfunktion des Schreibprogrammes (im Normalfall Word) zu erstellen; bitte nicht gesondert formatieren. Fußnoten sind nur für inhaltliche Kommentare vorzusehen, nicht für bibliographische Hinweise.

Literaturhinweise im Text sind nach den folgenden Mustern aufzuführen: Müller (2002) – Müller (2002: 75) – (vgl. Müller 2002: 75) – (Müller 2002; Mayer/Müller/Schulze 2003).

Das Literaturverzeichnis ist wie folgt zu gestalten:

Buchveröffentlichungen:

Strobl, R. und W. Kühnel, 2000: Dazugehörig und ausgegrenzt. Analysen zu Integrationschancen junger Aussiedler. Weinheim/München: Juventa.

Zeitschriftenbeiträge:

Becker, R., R. Imhof und G. Mehlkop, 2007: Die Wirkung monetärer Anreize auf den Rücklauf bei einer postalischen Befragung und die Antworten auf Fragen zur Delinquenz. Empirische Befunde eines Methodenexperiments. Methoden - Daten - Analysen. Zeitschrift für Empirische Sozialforschung 1 (2): 131-159.

Beiträge in Büchern:

Braun, M. und I. Borg, 2004: Berufswerte im zeitlichen und im Ost-West-Vergleich. S. 179-199 in: R. Schmitt-Beck, M. Wasmer und A. Koch (Hg.): Sozialer und politischer Wandel in Deutschland. Analysen mit ALLBUS-Daten aus zwei Jahrzehnten. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Internetquellen:

Stadtmüller, S. und R. Porst, 2005: Zum Einsatz von Incentives bei postalischen Befragungen. GESIS How-to-Reihe, Nr. 14 (Mannheim: GESIS). http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/howto/how-to14rp.pdf (1.12.2008).

